

Query No.	Score	Match	Length	DB	ID	Description
1	1828	100.0	351	21	AA850378	Human uncoupling P
2	1801	98.5	365	21	AA842783	Human OREF ORF2547
3	423.5	23.2	378	21	AA621416	Arbidopsin1 thallia
4	423.5	23.2	378	21	AA640152	Arbidopsin1 thallia
5	423.5	23.2	413	21	AA621415	Arbidopsin1 thallia
6	423.5	23.2	413	21	AA640151	Arbidopsin1 thallia
7	374	20.5	318	21	AA621417	Arbidopsin1 thallia
8	374	20.5	318	21	AA640153	Arbidopsin1 thallia
9	343	18.8	331	21	AA643092	Arbidopsin1 thallia
10	342	18.7	331	21	AA622077	Arbidopsin1 thallia
11	335	18.3	290	21	AA643093	Arbidopsin1 thallia

12	334	18.3	290	21	AA628078	Arabidopsis thaliana
13	323	17.7	278	21	AA633094	Arabidopsis thaliana
14	322	17.6	278	21	AA627079	Arabidopsis thaliana
15	320	17.5	289	18	AA617054	Blackcurrant RIB7
16	311	17.0	335	21	AA606050	Arabidopsis thaliana
17	299	16.4	305	21	AA606091	Arabidopsis thaliana
18	293	16.0	318	22	AA614505	Human polypeptide
19	286.5	15.7	300	21	AA613086	Arabidopsis thaliana
20	286.5	15.7	300	21	AA637830	Arabidopsis thaliana
21	286.5	15.7	300	21	AA639526	Arabidopsis thaliana
22	285	15.6	330	21	AA628434	Arabidopsis thaliana
23	285	15.6	332	21	AA628433	Arabidopsis thaliana
24	284.5	15.6	366	22	AA600938	Human bone marrow
25	284	15.5	459	21	AA628051	Arabidopsis thaliana
26	284	15.5	502	21	AA628050	Arabidopsis thaliana
27	284	15.5	617	21	AA628049	Arabidopsis thaliana
28	280.5	15.3	266	21	AA628435	Arabidopsis thaliana
29	280	15.3	289	22	AA650133	Human uncoupling P
30	280	15.3	289	22	AA650133	Human transport pr
31	271	14.8	305	22	AA640265	Human polypeptide
32	271	14.8	447	20	AAV05713	Human adenine nucle
33	271	14.8	674	20	AAV05712	Human adenine nucle
34	271	14.8	708	21	AA633247	Human OREF OREF3011
35	269.5	14.7	304	21	AA650384	Human uncoupling P
36	269.5	14.7	352	22	AA642051	Human polypeptide
37	258.5	14.1	357	22	AA630070	Arabidopsis thaliana
38	257	14.1	272	21	AA642980	Human OREF OREF2744
39	254	13.9	311	21	AA650382	Human uncoupling F
40	254	13.9	311	22	AA639173	Human polypeptide
41	251	13.7	323	21	AA633665	Arabidopsis thaliana
42	250.5	13.7	323	21	AA650419	Human uncoupling F
43	250	13.7	291	21	AA630071	Arabidopsis thaliana
44	247.5	13.5	266	22	AA639719	Human polypeptide
45	246.5	13.5	310	22	AA635001	Human protein sequ

ALIGNMENTS

RESULT 1

AAB50378 standard; Protein; 351 AA
AAB50378;

12-MAR-2001 (first entry)

Human uncoupling protein #1.

Human; uncoupling protein; immunosuppressive; antlarthritic;

[illegible]

gastrointestinal: nephrotropic; gynaeceological: vulnery: thrombolytic

gene therapy; cancer; wound; infectious disease; thrombosis; arthritis;

infertility.

Homo sapiens.

W0200061614-A2

19-OCT-2000.

06-APR-2000; 2000WO-US09534.

09-APR-1999; 99US-0128701.

18-AUG-1999; 99US-0149448.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Ruben SM, Ni J, Komatsoulis G, Rosen CA, Soppet DR;

DR WPI; 2000-656322/63.
 DR N-PSDB; AAC90452.
 XX
 PT Uncoupling proteins and nucleic acid sequences encoding them, useful
 PT for detecting, preventing and treating proliferative, neurological,
 PT immune system, cardiovascular and gastrointestinal disorders -
 XX
 XX
 PS Claim 11; Page 316-317; 343pp; English.
 CC The present sequence is a human uncoupling protein. The nucleotide
 CC sequences encoding the uncoupling proteins may be used for
 CC the detection of various disorders such as cancer, for chromosome
 CC identification, as chromosome markers and for numerous other diagnostic
 CC or research purposes. The uncoupling protein encoded by the nucleotide
 CC sequences may be used to treat disorders such as neural, immune,
 CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
 CC renal and proliferative disorders, wounds, infectious diseases,
 CC thrombosis, arthritis, and infertility.
 CC
 XX
 SQ Sequence 351 AA;

Query Match 100.0%; Score 1828; DB 21; Length 351;
 Best Local Similarity 100.0%; Pred. No. 1.3e-186;
 Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MADQDPAGISPLQOMWASGTGAVVTSLEMTPLDVVKVRLQSORPMSASELMPSSRLMSLS 60
 DB 1 madqdpagisplqgmwasgtgavvtslmtpldvkvrlqsgfpmaselmssrlmsls 60
 OY * 61 YTKWKCLLYCNGVLEPYLYCPNGARCATWQDPRTFTGMDAVKIVRHGRTIWSGLP 120
 DB 61 ytkwckllycngvlepylycpngarcatwqdprrtftgmdavkivrhgtrtlwsqilp 120
 OY 121 ATLMTVPATAIETFAVYDQKAFKALGSRALTSIDYAPVAVAGALARLGTWVISPLELMRTK 180
 DB 121 atlmtvpataiyftaydqkafkalgrraltsdlyapmvagalarlgtvrlsplelmrtk 180
 OY 181 LQAHQVSYRELACVRRPAAVAGQGRSLMLGWPALRDVPFSAIYWFNELYKSWMLNGLR 240
 DB 181 lqahvsvyrelagcvrrpaaavagqgrslmlgwpalrldvpfsaiylwfnelykswmlnglr 240
 OY 241 PKQGTSGVMSFVAGISGTYAAVLTLPDVVKTOROVAAICAMAAVVPNPLHVSSTWLLLR 300
 DB 241 pkqgtsvgmsfvaagisgtyaaavltpdvvktrdvaaicamaaavvpnpnhvstwlillr 300
 OY 301 RIRAESGTGKLFAGFLPRITIKAPSCAIMISYEFKSFQRLNDRLLAG 351
 DB 301 riraesgtgklfagflprilikaapscaimisyefksgffqrlndrlll 351

RESULT 2
 AAB42783
 ID AAB42783 standard; Protein; 365 AA.
 XX
 AC AAB42783;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human ORFX ORF2547 polypeptide sequence SEQ ID NO:5094.
 XX
 KW Human; open reading frame; ORFX; detection; cytosstatic; hepatotropic;
 KW vulnerability; antiparasitic; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritis; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antinflammatory;
 KW antiviral; antibacterial; antifungal; antineumatic; antithyroid;
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;

KW bone damage; cartilage damage; antinflammatory disease; coagulation;
 KW thrombosis; contraceptive.
 XX
 OS Homo sapiens.
 XX
 PN WO20058473-A2.
 XX
 PD 05-OCT-2000.
 XX
 PF 31-MAR-2000; 2000WO-US08621.
 XX
 PR 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shinkets RA, Leach M;
 XX
 DR WPI; 2000-602362/57.
 DR N-PSDB; AAC76992.
 XX
 PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 XX
 PS Claim 11; Page 4282-4283; 5507pp; English.

CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytosstatic; hepatotropic; vulnerability;
 CC antiparasitic; antiparkinsonian; nootropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritis; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antinflammatory; antibacterial; antiviral; antifungal; antineumatic;
 CC antithyroid; and antianemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 CC
 XX
 SQ Sequence 365 AA;

Query Match 98.5%; Score 1801; DB 21; Length 365;
 Best Local Similarity 96.2%; Pred. No. 1e-183;
 Matches 351; Conservative 0; Mismatches 0; Indels 14; Gaps 2;

OY 1 MADQDPAGISPLQOMWASGTGAVVTSLEMTPLDVVKVRLQSORPMSASELMPSSRLMSLS 57
 DB 1 madqdpagisplqgmwasgtgavvtslmtpldvkvrlqsgfpmaselmssrlmsls 60
 OY 58 -----SLSTYKWKCLLYCNGVLEPYLYCPNGARCATWQDPRTFTGMDAVKIVRHGRT 112
 DB 61 ytklpslytkwckllycngvlepylycpngarcattwqdprrtftgmdavkivrhgtr 120
 OY 113 RTLWSGLPATTMTVPATAIETFAVYDQKAFKALGSRALTSIDYAPVAVAGALARLGTWVVIS 172
 DB 121 rtlwsqilpattmtvpataiyftaydqkafkalgrraltsdlyapmvagalarlgtvrlvis 180
 OY 173 PLELMRTKIQAHQVSYRELACVRRPAAVAGQGRSLMLGWPALRDVPFSAIYWFNELYKSWML 226
 DB 181 plelmrtklqahvsvyrelagcvrrpaaavagqgrslmlgwpalrldvpfsaiylwfnelykswml 240

OY 227 FNYELKSWLNGLRPKDQTSVGSFVAGGISTVAVALTPDPVYKTOQROVALGAMEAVR 286
|||||
Db 241 fnyelkswlninglrpkdcgsvmsfvaegglsqtlvaavlllpddvktqfvalgameavr 300

OY 287 VNPLHWDSTWLLLRIRIRASGKGLFAGFLPRIRAKAFSCAIMISYERKSFQRLNOD 346
|||||
Db 301 vnplhwdstwlllrirraesgkglfagflprlrakafscaimisyltqksfqrlnqd 360

OY 347 RLUGG 351
|||||
Db 361 rlllg 365

RESULT 3
AAG21416
ID AAG21416 standard; Protein: 378 AA.
XX
AC AAG21416;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 23965.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142184.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 22-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147182.
PR 05-AUG-1999; 99US-0147260.

PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140822.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142380.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148177.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159284.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.

	PR	21-OCT-1999;	99US-0160815.
	PR	22-OCT-1999;	99US-0160980.
	PR	22-OCT-1999;	99US-0160981.
	PR	22-OCT-1999;	99US-0160989.
	PR	25-OCT-1999;	99US-0161404.
	PR	25-OCT-1999;	99US-0161405.
	PR	25-OCT-1999;	99US-0161406.
	PR	26-OCT-1999;	99US-0161359.
	PR	26-OCT-1999;	99US-0161360.
	PR	26-OCT-1999;	99US-0161361.
	PR	28-OCT-1999;	99US-0161920.
	PR	28-OCT-1999;	99US-0161992.
	PR	28-OCT-1999;	99US-0161993.
	PR	29-OCT-1999;	99US-0162142.
	Query Match	23.2%;	Score 423.5; DB 21; Length 378;
	Best Local Similarity	32.0%;	Pred. No. 14e-36;
	Matches 117;	Conservative 57;	Mismatches 137; Indels 55; Gaps 9
OY	13	OOMVASGCAVVTSLFMTPLDVVKYRLQSORPSMASLMPSSRLM-----SLSYTKMK	65
Dd	28	evfvaagaavivsevtlpldvtkrtllqgaagmayshtplsmsigymafgpmmtfadlr	87
OY	66	CILLYC--NGVLEPPLYLCENGCARCATWFODPTFTGTMDAFVKIVRHEGTRITLMSELPALU	123
Dd	88	cspscaraegvegtvslcp-----pcdfykgflfdvfktllrqdgilgrwtgnagl	138
OY	124	VMTVEPATVIIFYPAVD---QLKAPFLCGRALITSDLVAPMVAGALARIGTVTVISPLELMT	179
Dd	139	alayrmvgylprfydmfnmrllcelesrekapantfcvptvagstlaractocypidiatl	198
OY	180	KIQAOHVSTRDELGAC-----VRTA----VAQGWRSLMLGMGPTRLRDV	219
Dd	199	rmgqa-----fkkeakgvkppgvftklvygvisevrytanmllesslhnyrgtlwrglgaqlardv	254
OY	220	PRSALTWMFYETELVKSGLNLRPKDDQSVMG---SFVAGGISGVAAVLTLPEPDVKTORQ	276
Dd	255	pstaicwsletpikrlllgvagnndnlvgatfsagfiagsiaaaacopldvatrrtg	314
OY	277	VALGAMEAVRVNPDLHVDSTWLLRLIRIAESGTKGLFAGFLPRIIKAPSICALMIISTVERG	336
Dd	315	lekdpgrlatmm-----ttrqtllievwrsgmrglimgmpravaragpsvglavsfigevv	368
OY	337	KSEFPQR	342
Dd	369	kylvlnr	374
RESULT	5		
ID	AAG21415		
XX	AAG21415 standard:	Protein; 413 AA.	
XX	AAG21415;		
DE	17-OCT-2000	(first entry)	
XX	Arabidopsis thaliana protein fragment SEQ ID NO: 23964.		
KW	Protein identification; signal transduction pathway; metabolic pathway;		
KW	hybridisation assay; genetic mapping; gene expression control; promoter;		
OS	terminalin sequence.		
XX	Arabiidopsis thallana.		
PN	EP1033405-A2.		
PD	06-SEP-2000.		
PF	25-FEB-2000;	2000EP-0301439.	
XX	25-FEB-1999;	99US-0121825.	
PR	05-MAR-1999;	99US-0123180.	
PR	09-MAR-1999;	99US-01233548.	

PR	25-MAR-1999;	9905-0126264
PR	25-MAR-1999;	9905-0126788
PR	29-MAR-1999;	9905-0128785
PR	01-APR-1999;	9905-0127462
PR	06-APR-1999;	9905-0128234
PR	08-APR-1999;	9905-0128714
PR	16-APR-1999;	9905-0128845
PR	19-APR-1999;	9905-0130077
PR	21-APR-1999;	9905-0130449
PR	23-APR-1999;	9905-0130510
PR	23-APR-1999;	9905-0130891
PR	28-APR-1999;	9905-0131449
PR	30-APR-1999;	9905-0132048
PR	30-APR-1999;	9905-0133407
PR	04-MAY-1999;	9905-0133484
PR	05-MAY-1999;	9905-0133485
PR	06-MAY-1999;	9905-0133486
PR	06-MAY-1999;	9905-0133487
PR	07-MAY-1999;	9905-0132863
PR	11-MAY-1999;	9905-0134216
PR	14-MAY-1999;	9905-0134218
PR	14-MAY-1999;	9905-0134219
PR	14-MAY-1999;	9905-0134221
PR	14-MAY-1999;	9905-0134370
PR	14-MAY-1999;	9905-0133768
PR	18-MAY-1999;	9905-0133941
PR	19-MAY-1999;	9905-0135124
PR	20-MAY-1999;	9905-0135153
PR	21-MAY-1999;	9905-0135253
PR	24-MAY-1999;	9905-0136021
PR	25-MAY-1999;	9905-0136029
PR	27-MAY-1999;	9905-0136392
PR	28-MAY-1999;	9905-0136788
PR	01-JUN-1999;	9905-0137222
PR	04-JUN-1999;	9905-0137502
PR	04-JUN-1999;	9905-0137528
PR	07-JUN-1999;	9905-0137724
PR	08-JUN-1999;	9905-0138094
PR	10-JUN-1999;	9905-0138547
PR	10-JUN-1999;	9905-0138840
PR	14-JUN-1999;	9905-0139311
PR	16-JUN-1999;	9905-0139451
PR	16-JUN-1999;	9905-0139453
PR	17-JUN-1999;	9905-0139492
PR	18-JUN-1999;	9905-0139454
PR	18-JUN-1999;	9905-0139455
PR	18-JUN-1999;	9905-0139456
PR	18-JUN-1999;	9905-0139457
PR	18-JUN-1999;	9905-0139458
PR	18-JUN-1999;	9905-0139459
PR	18-JUN-1999;	9905-0139460
PR	18-JUN-1999;	9905-0139461
PR	18-JUN-1999;	9905-0139462
PR	18-JUN-1999;	9905-0139463
PR	18-JUN-1999;	9905-0139750
PR	18-JUN-1999;	9905-0139763
PR	21-JUN-1999;	9905-0139817
PR	22-JUN-1999;	9905-0139899
PR	22-JUN-1999;	9905-0140353
PR	23-JUN-1999;	9905-0140354
PR	24-JUN-1999;	9905-0140695
PR	28-JUN-1999;	9905-0140823
PR	29-JUN-1999;	9905-0140991
PR	30-JUN-1999;	9905-0141287
PR	01-JUL-1999;	9905-0141842
PR	01-JUL-1999;	9905-0142154
PR	02-JUL-1999;	9905-0142055
PR	06-JUL-1999;	9905-0142930
PR	08-JUL-1999;	9905-0142803
PR	09-JUL-1999;	9905-0142920
PR	12-JUL-1999;	9905-0142977
PR	13-JUL-1999;	9905-0143542
PR	14-JUL-1999;	9905-0143624
PR	15-JUL-1999;	9905-0144005

DE Arabidopsis thaliana protein fragment SEQ ID NO: 49781.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KM hydrolisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134321.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143544.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.

```
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154739.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.
```

Query Match 23.2%; Score 423.5; DB 21; Length 413;

Best Local Similarity 32.0%; Pred. No. 1.6e-36; Matches 117; Conservative 57; Mismatches 137; Indels 55; Gaps 9;

```
OY 13 QQMYASTGAVTSLFMTPLDVYKVRLOQSPMSASLMPSSRRM-----SLSTYKWK 65
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 63 ervfsaagaavlsavtlnpdlvkrtrlqagaagmsysbplsnstgirmaffgpmmmiadlr 122
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 66 CLAYC--NGVLEPLYLCPNGARCACTWFODPTRFGTMDAVKTIYRHGCTRTMNSGLPATL 123
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 123 cspcacaagvgtvslcp-----pdcfkykgtldvrfkllirgsglgrlwtgtaagl 173
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 124 VMTVPATAIYFTAYD---OLKAFLCGRALTSDLYAPMVAGALRLGTVTVISPLELMRT 179
      : : | | | | : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 174 alavpmwgiylpfdmfrnlleelsrckapamfcvptvagslarslactvcypidart 233
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 180 KLOQOHYSRELKAC-----VRTA-----VAQGWBSLMLGMPFALRDV 219
      : : | | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 234 rmga-----fkeakagvppgyfkllygvfsevrtaannleslhnyrglwtgltgaqlardv 289
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 220 PFSALYWFNVELYKSWMLNGLRPKDOTSYGM---SFVAGGISGTVAAVLLTFEDVVKTORQ 276
      | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 290 pfsalcwstleplkrlilgvgndtnllyvgvfatfsagflagslaaatactopldvarcrrg 349
```

```
OY 277 VALGAEAVRNPLHVDSTWLLLRIRAESGTGKGLPAGFLPPIITKAAPSALMISTYERG 336
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 350 iekdpgralmm-----trrgtliewrdgmgtrglfmgmgpvaragsvgtvafygev 403
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 337 KSFFQR 342
      | | | | |
DB 404 kylvlr 409
      : : : : :

RESULT 7
AAG21417
ID AAG21417 standard; Protein; 318 AA.
XX
AC AAG21417;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 23966.
XX
KW 'Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
XX
PR 05-MAR-1999; 99US-0123180.
XX
PR 09-MAR-1999; 99US-0123548.
XX
PR 23-MAR-1999; 99US-0125788.
XX
PR 25-MAR-1999; 99US-0126264.
XX
PR 29-MAR-1999; 99US-0126785.
XX
PR 01-APR-1999; 99US-0127462.
XX
PR 06-APR-1999; 99US-0128234.
XX
PR 08-APR-1999; 99US-0128714.
XX
PR 16-APR-1999; 99US-0129845.
XX
PR 19-APR-1999; 99US-0130077.
XX
PR 21-APR-1999; 99US-0130449.
XX
PR 23-APR-1999; 99US-0130510.
XX
PR 28-APR-1999; 99US-0130891.
XX
PR 30-APR-1999; 99US-0131449.
XX
PR 04-MAY-1999; 99US-0132048.
XX
PR 05-MAY-1999; 99US-0132484.
XX
PR 06-MAY-1999; 99US-0132485.
XX
PR 07-MAY-1999; 99US-0132863.
XX
PR 11-MAY-1999; 99US-0134256.
XX
PR 14-MAY-1999; 99US-0134218.
XX
PR 14-MAY-1999; 99US-0134219.
XX
PR 14-MAY-1999; 99US-0134221.
XX
PR 14-MAY-1999; 99US-0134370.
XX
PR 18-MAY-1999; 99US-0134768.
XX
PR 19-MAY-1999; 99US-0134941.
XX
PR 20-MAY-1999; 99US-0135124.
XX
PR 21-MAY-1999; 99US-0135353.
XX
PR 24-MAY-1999; 99US-0135629.
XX
PR 25-MAY-1999; 99US-0136021.
XX
PR 27-MAY-1999; 99US-0136392.
XX
PR 28-MAY-1999; 99US-0136782.
XX
PR 01-JUN-1999; 99US-0137222.
XX
PR 03-JUN-1999; 99US-0137528.
XX
PR 04-JUN-1999; 99US-0137502.
XX
PR 07-JUN-1999; 99US-0137724.
XX
PR 08-JUN-1999; 99US-0138094.
XX
PR 10-JUN-1999; 99US-0138540.
XX
PR 10-JUN-1999; 99US-0138847.
```

```
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 29-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140991.
PR 01-JUL-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.

PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149829.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159684.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160747.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161040.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 29-OCT-1999; 99US-0162142.

Query Match 20.5%; Score 374; DB 21; Length 318;
Best Local Similarity 32.8%; Pred. No. 2.2e-31;
Matches 103; Conservative 43; Mismatches 120; Indels 48; Gaps 8;

QY 58 SLSYTKWKCLAYC--NGVLEPLYLCPNGARCATWFODFRFRFGTMDAVKIVRHGCTRTL 115
DB 20 nmmtaditrcpscargavgevtwscp-----pdcdfykygltvdvltkiltirgeglgrl 70
```



```

PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147935.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 23-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.

```

```

PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 20.8%; Score 374; DB 21; Length 318;
Best Local Similarity 32.8%; Pred. No. 2, 2e-31;
Matches 103; Conservative 43; Mismatches 120; Indels 48; Gaps 8;

OY 58 SLSTYKWKCLLYC--NGVLEPLYLCPNGARCATWFODPRTFTGMDAFYKIVRHGSTRLL 115
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 20 nmffadrlrcpsccaragvegtvsicp-----pdcftgyktfdvfkllrgeqlrgl 70

OY 116 WSGLPATLWTVPAATAYTFAVD---OLKAPFLCGSRLSDLYAPVAGALRLGTVPTI 171
   | | | | : : | | | | : : | | : : | | | | | | | | | | | | | | | |
DB 71 wrtgtnaglalavpwngvlylptfdmfrlnlejsrekapamtfcvplvagslarslactvc 130

OY 172 SPLELMRTKLQAOHVSRYRELGAC-----VPTA---VAQGGMRSLMLGM 211
   | : | | : | | : : : : : : : : : : : : : : : : : : : : : : : :
DB 131 ypldlatrtmga-----fkeakagvdkppgyfktlvgvfsevrftannlesslmyrqlwrgl 186

OY 212 GPTALBDVPFSALYFNTLVLKSWLNGLRPKDQTSVGM--SPVAGGISGTVAAVLTLPF 268
   | : | | | | | : : : : : : : : : : : : : : : : : : : : : : : :
DB 187 gqglardvypfsalcwstlepkikrllvgandtnlvgvfatfsagflagslaaatcpl 246

OY 269 DVVKTORQVALGAMEAVRNPLHVDSTWLLERIRAESGTRKGLFNGFLPRITIKAPSCAI 328
   | | : | | : | : : : : : : : : : : : : : : : : : : : : : : :
DB 247 dvartirgkdkpgralnm-----ltrqtlievrdsgmrglimgmprvaragsvyl 300

OY 329 MISTYEFGRKSPFOR 342
   : : | | | | | |
DB 301 vvsfyevwkyvlhr 314

RESULT 9
AAG43092
ID AAG43092 standard; Protein; 331 AA.
XX
XX AAG43092;
AC
XX
XX 18-OCT-2000 (first entry)
DT
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 53820.
DE
XX
XX Protein identification: signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX
XX EP1033405-A2.
FN
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-0301439.
PE
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.

```


PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 23-JUN-1999; 99US-0140355.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141847.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144633.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0144815.
PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147933.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149922.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151067.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0155648.
PR 29-SEP-1999; 99US-0155696.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.

PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 18.8%; Score 343; DB 21; Length 331;

Best Local Similarity 27.8%; Pred. No. 4.7e-28;

Matches 99; Conservative 58; Mismatches 133; Indels 66; Gaps 7;

OY 3 DDPAGISPLDQ-----MVASGTGAVVTSLEMTPLDVKVRLOQSRPSMASSEIMPSS 54.
DB 22 dhpaiivpaquttlkfmglmwagslagsvehmamfpvdtvkthmgalr----- 70
OY 55 RLMSLSYRKWKCLXCNVLEPLVLCPPNGARCATWFODPTRFTGMDARVKIVRHGRTT 114
DB 71 -----sc-----plkpiqirgafrslkiktdgsa 94
OY 115 LMSGLPATLVMTVPATAYFTAYDOLKAFLCGRALTSIDYAPMVAGALARLGTVTVISPL 174
DB 95 lyrgtwanglgaqphavayfsfeyevskkfllsg-gpnmsaahaisgfvfctissdavidfpm 153
OY 175 ELMRKLDQAHVSYRELGCACVNTAVAGGWSLWLGMEFTALRDVPFSALYFNTELYKS 234
DB 154 dmvkqrlqigngtykvavdciktrveegfigafyasyltvlmnapflavhlttyeaavr 213
OY 235 WLNGLRP-----KDQTSVGSMSFVAGISGTVAAVLTLPFDVVKTOROVALGAMEAVERN 288
DB 214 glremhpehayaadeegwilyatagaagajlaaavtupldvckcqldq-cgvgcgcdrtk 272
OY 289 PLHVDSTWLLMRIRAESGTGKLFAGFLPRIRIKAPSCAIMISTVEFGKSPFORLN 344
DB 273 sssisd---vftliykkdgyrglaryglpmlfhapaaicwstyetvskffqldn 325

RESULT 10

AAG22077
ID AAG22077 standard; Protein: 331 AA.

AC AAG22077;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 24869.

KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123160.
PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0125264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134376.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.

termination sequence.
XX OS Arabidopsis thaliana.
XX PN EPI03405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138099.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145224.
PR 23-JUL-1999; 99US-0145226.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146389.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148365.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.

PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139819.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145199.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148564.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.

PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150584.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158222.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 18.3%; Score 334; DB 21; Length 290;

Best Local Similarity 27.7%; Pred. No. 3, 6e-27; Mismatches 127; Indels 58; Gaps 6;

Matches 93; Conservative 58; Mismatches 127; Indels 58; Gaps 6;

QY 15 MWASGTGAVVTSLEFMPDLVVKVRLQSRPSMASSELMPPSSRLMSLSTKWKCLLYNGVYL 74
DB 1 mvaagslsgsvvhamfvdvctklmgqlr-----sc----- 31

QY 75 EPLYLCPNGARCATWFPQDPFRFTGTMDFVKIYRHEGTRTMSGLPATLVMTVPATATPF 134
DB 32 -----plkpiqitqratslilktdpsalslygiwamlgagpahavyf 73

QY 135 TAYDOLKAPFLGGRALTSYDIAFVAPVACALARLGTIVTISPLELRTKIQAOHVSREILGAC 194
DB 74 sfyevskkfllsg--gnpnnsahalsgvyfatlsdavgftpmamvkgqlqngnfykgvwdc 132

PR	06-AUG-1999;	99US-0147203.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0153930.
PR	07-SEP-1999;	99US-0153263.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157863.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	18-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	22-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161993.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.
Query Match		
Best Local Similarity	17.7%;	Score 323; DB 21; Length 278;
	27.6%;	Pred. No. 5,1e-26;

Matches	89; Conservative	57; Mismatches	118; Indels	58; Gaps	6;
QY	29	MPLDVKVRLQSQRPSMASSELPSSRLMSLSYTKWKCLLYCNGVLEPLYPNGRCAT			
DB	3	mfpvdtvktchmqalr-----sc-----			
QY	89	WFODDPRTFTGMDAFKVIKRIHREGTRIMSGLPATLIMVIPAATAIYETAVDOLKAEICGRA			
DB	20	----pkpigiqlqatrsilktqpsalYrqlwamglgagpahavYfsfYevskklts-g			
QY	149	LTSIDYAPMVAGALARLGTIVISPLEMRTKIQAOHVSRELGACVPTAAVAGGMSRLW			
DB	75	npnusaahaisgYfatissdavftpmdwkqrlqYngntlykYvdcikrvtreegfgafY			
QY	209	LWMGPTALRDVFPFSALYWNVELYKSWLNGLRP-----KDOTSVGMSFVAGISGTVAA			
DB	135	asYrtivlmnapflavhltYeaYkrqlremIpehavaadeegwlYatagaaagqlaa			
QY	263	VLTLPEDVVKTRQVALGMEAVRVNPLHVDSTWLLRRIRAESGTGKLFAGFLPRIKA			
DB	195	avtlpdlavvktqlg-cqYvcgcdrfkssissd---vftivYkdgYrqlarYwlpmlfh			
QY	323	APSCAIMISTYEFKGSFFORLN			
DB	251	apaaalcwstYelvtsfifqdlm			
RESULT 14					
ID	AAG22079 standard; Protein: 278 AA.				
XX	AAG22079;				
DT	17-OCT-2000 (first entry)				
XX	Arabidopsis thaliana protein fragment seq ID NO: 24871.				
XX	Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.				
XX	Arabidopsis thaliana.				
OS	EP1033405-A2.				
PN	EP1033405-A2.				
XX	06-SEP-2000.				
PD	06-SEP-2000.				
XX	25-FEB-2000; 2000EP-0301439.				
PF	25-FEB-2000; 2000EP-0301439.				
XX	25-FEB-1999; 99US-0121825.				
PR	05-MAR-1999;	99US-0123180.			
PR	09-MAR-1999;	99US-0123548.			
PR	23-MAR-1999;	99US-0125788.			
PR	25-MAR-1999;	99US-0126264.			
PR	29-MAR-1999;	99US-0126785.			
PR	01-APR-1999;	99US-0127462.			
PR	06-APR-1999;	99US-0128234.			
PR	16-APR-1999;	99US-0128714.			
PR	19-APR-1999;	99US-0129845.			
PR	21-APR-1999;	99US-0130077.			
PR	23-APR-1999;	99US-0130449.			
PR	23-APR-1999;	99US-0130510.			
PR	23-APR-1999;	99US-0130891.			
PR	28-APR-1999;	99US-0131449.			
PR	30-APR-1999;	99US-0132048.			
PR	30-APR-1999;	99US-0132407.			
PR	04-MAY-1999;	99US-0132484.			
PR	05-MAY-1999;	99US-0132485.			
PR	06-MAY-1999;	99US-0132486.			
PR	06-MAY-1999;	99US-0132487.			
PR	07-MAY-1999;	99US-0132863.			
PR	11-MAY-1999;	99US-0134256.			
PR	14-MAY-1999;	99US-0134218.			

PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134321.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134766.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140699.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142153.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144634.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147436.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156556.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2002, 21:47:27 ; Search time 69.69 Seconds
(without alignments)
383.660 Million cell updates/sec

Title: US-09-840-787-19

Perfect score: 1828
Sequence: 1 MADDPAGISPLQOMVASGT.....TYEFGKSFFORLNDRLG 351

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

PIR_60: *
1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1828	100.0	351	2 T43493	hypothetical prote
2	488.5	35.5	328	2 T19322	hypothetical prote
3	486.5	26.6	335	2 T50393	probable mitochond
4	474.5	26.0	366	2 S64589	hypothetical prote
5	348	19.0	358	2 D84901	hypothetical prote
6	343	18.8	331	2 T00582	probable mitochond
7	312	17.1	781	2 A86305	hypothetical prote
8	300.5	16.4	378	2 T02895	hypothetical prote
9	286	15.6	332	2 T47703	Ca-dependent solut
10	284.5	15.6	902	1 S54495	probable carrier p
11	281.5	15.4	310	2 S69050	probable membrane
12	277.5	15.2	348	2 D84798	probable mitochond
13	273.5	15.0	358	2 T45934	hypothetical prote
14	272	14.9	312	2 T25089	hypothetical prote
15	268.5	14.7	368	2 S54524	hypothetical prote
16	264	14.4	702	2 T16533	hypothetical prote
17	260.5	14.3	302	2 T38879	probable mitochond
18	259.5	14.3	338	2 T40968	probable phosphate
19	255.5	14.2	307	2 S60949	hypothetical prote
20	255.5	14.2	309	2 T29235	mitochondrial solu
21	258.5	14.1	352	2 T01729	mitochondrial solu
22	257	14.1	475	2 T50686	peroxisomal Ca-dep
23	256	14.0	303	2 T39149	probable RNA splic
24	254	13.9	349	2 B96753	hypothetical prote
25	250.5	13.7	309	2 T48156	hypothetical prote
26	249.5	13.6	325	2 T04273	hypothetical prote
27	249.5	13.6	330	2 T29640	mitochondrial carr
28	249	13.6	304	2 S13533	mRNA splice defec
29	245.5	13.4	377	2 S36081	probable carrier p

30	243.5	13.3	322	2 S57116	probable carrier p
31	237	13.0	347	2 T50990	hypothetical prote
32	235	12.9	343	2 T15253	hypothetical prote
33	233	12.7	330	2 S26596	Graves disease mit
34	233	12.7	373	2 S48451	probable membrane
35	232	12.7	314	2 S51179	mRNA splice defec
36	229	12.5	327	2 S61660	probable mitochond
37	228.5	12.5	447	2 T00435	probable membrane
38	227.5	12.4	339	2 A41677	ADP-ATP carrier pr
39	226.5	12.4	284	2 S45458	uncoupling protein
40	226.5	12.4	313	2 T05577	probable mitochond
41	223	12.2	345	2 T37576	hypothetical prote
42	223	12.2	415	2 T48171	Btl protein precu
43	221	12.1	436	2 J01459	probable membrane
44	220	12.0	326	2 S57544	probable membrane
45	219.5	12.0	357	2 S46795	hypothetical prote

ALIGNMENTS

RESULT 1

T43493
hypothetical protein DKFzp434C119.1 - human

C:Species: Homo sapiens (man)

C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000

C:Accession: T43493

R:Poustka, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, December 1999

A:Reference number: 22516

A:Accession: T43493

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-351 <AAA>

A:Cross-reference: EMBL:AL133584

A:Experimental source: adult testis; clone DKFzp434C119

C:Genetics:

A>Note: DKFzp434C119.1

Query Match	100.0%	Score 1828;	DB 2;	Length 351;
Best Local Similarity	100.0%	Pred. No. 3.1e-149;		
Matches 351;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MADDPAGISPLQOMVASGTGAVVTSLEFMPDLVYKVKLOSQSPMSASELPPSSRLMSLS	60	
DB	1	MADDPAGISPLQOMVASGTGAVVTSLEFMPDLVYKVKLOSQSPMSASELPPSSRLMSLS	60	
QY	61	YTKMKCLYCNGLVLEPLVLCNGARCATWPDPTREFTGTMDAFVKIYRHEGTRTLMGSLP	120	
DB	61	YTKMKCLYCNGLVLEPLVLCNGARCATWPDPTREFTGTMDAFVKIYRHEGTRTLMGSLP	120	
QY	121	ATLVMTVPATAYFTAYDOLKAFGLGRALTSDLVAPMVAGALRLGTVTVISPELMRTK	180	
DB	121	ATLVMTVPATAYFTAYDOLKAFGLGRALTSDLVAPMVAGALRLGTVTVISPELMRTK	180	
QY	181	LOAQHVSRELCACVTRTAAVAGGMSLWLGNCPTLRVPVSALYWFYELVYKSWLNLGR	240	
DB	181	LOAQHVSRELCACVTRTAAVAGGMSLWLGNCPTLRVPVSALYWFYELVYKSWLNLGR	240	
QY	241	PKDQTSVGSFVAGISGTVAAVLTLPEDVYKTRQVALGMAEAVRNPVLHVDSTWLLLR	300	
DB	241	PKDQTSVGSFVAGISGTVAAVLTLPEDVYKTRQVALGMAEAVRNPVLHVDSTWLLLR	300	
QY	301	RIRAESGTGLFAGFLPRITKAAPSCAIIWISYEFEGKSFFORLNDRLG 351		
DB	301	RIRAESGTGLFAGFLPRITKAAPSCAIIWISYEFEGKSFFORLNDRLG 351		
RESULT 2				
T19322		hypothetical protein C16C10.1 - Caenorhabditis elegans		
C:Species: Caenorhabditis elegans				

Db 332 YTGIAARVYIKIRPSCAIISSYEISKVFGNKLHQ 366

RESULT 5

Db4901

hypothetical protein At2g46320 (imported) - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: DB4901

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanhaken, S.E.; Umayam, L.; Tallon, L.; eus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: AB4420; MUID:20083487

A:Accession: DB4901

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-358 <STO>

A:Cross-references: GB:AE002093; NID:96598574; PIDN:AAFL18629.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g46320

A:Map position: 2

Query Match 19.0%; Score 348; DB 2; Length 358;

Best Local Similarity 28.1%; Pred. No. 2,8e-22;

Matches 103; Conservative 61; Mismatches 137; Indels 66; Gaps 10;

QY 13 QOMVASTGAVVTSLEFMPPLDVYKVRLOSGRPSMASSELMSSRLMSLSTKKKCLLYC-- 70

Db 23 ERLASAGGAFTSAVYVNPDLVAKTRLOAGAVPYQ-----GSCRIGCPD 68

QY 71 -----NGVLEPLTLCNGARCATWFDPRFRGTMDAFYKIVRHEGRTLMGSL 119

Db 69 TNSTLVHDIRNSAPGMCRTISATYCS-----DNQKGTLLVFYKTLTIOSTGT----- 118

QY 120 PATL-----VMTPTATVAYFYTAVDQ-----LKAFLCGRALTSIDLAPVAGALARTGVYI 171

Db 119 --TLCKKGLGFGVGIYVPCYDFRNIMEFTTEKSPSLTYVYVPLVAGTIRSLACISC 176

QY 172 SPLELMRTKLOAHVSYR--ELGACVRYAV-----AOGGMSLWLGMPALRDVPF 221

Db 177 YVELLARTKQAFKGTQRNVKTLPGVWKTLDVYVNPYKGSNNQYRMLMTGLGAQLARDVPF 236

QY 222 SALVWNPVLYVSKWN---GLRPKQOTSYGMSFVAGISGTAAVLTLPEDVVKTOROVA 278

Db 237 SAICWSTLEPRTSRISQISANGEEPRAGSTIGANFAGFVAGVAAVAAATCPLDVAKTRQTE 296

QY 279 LGAMEAVRNPLHVSTWLLRIRAESGTLGFLPRTIKAAPSCAIISSYEFGKS 338

Db 297 KMTDRAMFM-----TTROTIAEIMRDGMRGMSGAGARVAGSVAIVVSFYEYVKY 350

QY 339 FFORLNO 345

Db 351 GLHNFHQ 357

RESULT 6

T00582

probable mitochondrial carrier protein (imported) - Arabidopsis thaliana

N:Alternate names: hypothetical protein T27E13.10

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 23-Mar-2001

C:Accession: T00582; C64705

R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul, M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanhaken, S.E.; Umayam, L.; Tallon, L.; eus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

submitted to the EMBL Data Library, May 1998

A:Description: Arabidopsis thaliana chromosome II BAC T27E13 genomic sequence.

A:Reference number: 214178

A:Accession: T00582

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-331 <ROU>

A:Cross-references: EMBL:AC004165; NID:93150396; PID:93150404

A:Experimental source: cultivar Columbia

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y. M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanhaken, S.E.; Umayam, L.; Tallon, L.; eus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: AB4420; MUID:20083487

A:Accession: C64705

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-331 <STO>

A:Cross-references: GB:AE002093; NID:93150404; PIDN:AAFL18629.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g30160; T27E13.10

A:Map position: 2

A:Introns: 263/3

C:Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology

C:Keywords: mitochondrion

Query Match 18.8%; Score 343; DB 2; Length 331;

Best Local Similarity 27.8%; Pred. No. 6,8e-22;

Matches 99; Conservative 58; Mismatches 133; Indels 66; Gaps 7;

QY 3 DDDPAGISPLDQ-----MVASGTGAVVTSLEFMPPLDVYKVRLOSGRPSMASSELMSS 54

Db 22 DHPALIVPAQNTTLKFMQWVAGSIAGSVHEHMAFEPDVTYKTHQALR----- 70

QY 55 RLMSLSYTKMKCLYCNGLVLEPLTLCPNGARCATWFDPRFRGTMDAFYKIVRHEGRT 114

Db 71 -----SC-----PIKIGIGROAFRSIKTKDGSA 94

QY 115 LMSGPLATLVMTVPATVAYFYTAVDQLKAFGLCGRALTSIDLAPVAGALARTGVISPL 174

Db 95 LRGVWAGLGGPRHAYFTSYEVSKFLSG-GNPNSAAIAISGVATISSDAVFTPM 153

QY 175 ELMTKLOAHVSYRELGAQVTRAVAAOGGMSLWLGMPALRDVPFSAIYFNTELYKS 234

Db 154 DMVKQRLOQNGCYGVMDICIRVTRREGFGAFVASYRTVLMNAPFVAFHTYEAVYR 213

QY 235 WINGLRP-----KQOTSYSMSFVAGISGTAAVLTLPEDVVKTOROVALGAMENVRYN 288

Db 214 GIREMPLBEHVAAGEDEGMLIYATGAAAGLAATPPLDVVKTKQLQ-CQGVCGCDRKK 272

QY 289 PLHVSTWLLRIRAESGTLGFLPRTIKAAPSCAIISSYEFGKSFORLN 344

Db 273 SSSISD--VFRTYVKKDQKGLARGLPRLMFLHAPAAIICMSTYETVYKSPFDOLN 325

RESULT 7

AB6205

hypothetical protein (imported) - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: AB6205

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon, Chiu, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewart, ansen, N.F.; Hughes, B.; Hufar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Matel, R.; Matzila, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: AB6141; MUID:21016719

A:Accession: AB6205

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-781 <STO>

A:Cross-references: GB:AE005172; NID:96954043; PIDN:AAFL2217.1; GSPDB:GN00141

C:Genetics:

A:Gene: MIPS:YPR021c
A:Map position: 16R
C:Superfamily: probable carrier protein YPR021c: ADP/ATP carrier protein repeat homology
C:Keywords: duplication, transmembrane protein
F:527-615/Domain: ADP/ATP carrier protein repeat homology <ACP1>
F:621-711/Domain: ADP/ATP carrier protein repeat homology <ACP2>
F:724-814/Domain: ADP/ATP carrier protein repeat homology <ACP3>

RESULT 13

T45934
hypothetical protein F5K20_240 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 17-Mar-2000
C:Accession: T45934
R:Konfort, A.; Casachubeta, E.; Pulgomech, P.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X
submitted to the Protein Sequence Database, January 2000
A:Reference number: Z23017
A:Accession: T45934
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-358 <MON>
A:Cross-references: EMBL:AL132960
A:Experimental source: cultivar Columbia; BAC clone F5K20
C:Genetics:
A:Map position: 3
A:Introns: 98/3; 157/3; 207/3; 234/3; 258/2; 281/1
A>Note: F5K20_240
A:Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology

Query Match	15.0%	Score 273.5;	DB 2	Length 358;
Best Local Similarity	24.7%	Pred. No. 6.9e-16;		
Matches 84; Conservative	50;	Mismatches 145;	Indels 61;	Gaps 6

```

QY 3 DODPAGISPLDOOMVASTGVAWVLSLFMTPLDDVKKVLRLOSQRPSMASLSSRLMSLSY 6
Db 61 NOQGGHGTVERILLAGIAGAFSKTCYAPLARLILFOLIOGMOSEALILSSPIMTHEA-- 118
QY 63 KWKOLLICNGLEBYLUCPNCARATCFODPRTRETCMDAFVILVHREGRITLMSGLPAT 122
Db 119 -----SRVKEEGRAWKGNLYT 137
QY 123 LVMTVPATAYETAYDOLKAFICLRALTS-----DLVAPVAGALARI GTVTVISPL 174
Db 138 VAHRLPGAVHFVYEEKTFELHNPLOSQYKNAGVADISVHNEFSGSLAGLTFAASATYPL 197
QY 175 ELMTSTYLOAOHVSYRELGACVETPAAGGMYSLMLGSPALRDVPSALUYWNYELKS 234
Db 198 DLVETRLSAQGV-----GHAFRTICREBGILIGLKGLGATLLGVGSPALISPAAYETFKT 252
QY 235 WLNGRLRPDQTSVGMSPFVAGISGTVAAVLTLPEDVYKTOQOVALGAMEAVRNPLHVD 294
Db 253 FWLSHRPDMSNAV--VSLGCGSLSGIVSSTAFTEPLDLVRRMQLE-GAGCARAVYTTGLFG 310
QY 295 TWLLLRIRAESGTRKGLFAGLPLRIIRAAAPSCAIMISTYE 334
Db 311 TF-----KHLEFTEGMGLRGILIPYKVVAVPVGIAEMTFE 347

```

RESULT 14

T26089
hypothetical protein W02B12.9 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
C:Accession: T26089
R:S.Windburne, J.: Ainscough, R.
submitted to the EMBL data library, October 1995
A:Reference number: Z20149
A:Accession: T26089
A:Status: preliminary; translated from GR/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-312 <M1>
A:Cross-references: EMBL:Z66521; PIDN:CAA91399.1; GSPDB:GN00020; CESP:W02B12.9
A:Experimental source: clone W02B12
C:Genetics:
A:Gene: CESP:W02B12.9
A:Map position: 2
A:Introns: 18/3; 251/3; 286/3
C:Superfamily: ADP/ATP carrier protein; ADP/ATP carrier protein repeat homology

Query Match:	14.9%	Score 272:	DB 2:	Length 312:
Best Local Similarity:	24.6%	Pred. No. 7:	8-16:	
Matches	82:	Conservative	54:	Mismatches 134: Indels 64: Gaps

QY	15	MVASGTAAVYTSLEMPDLVDYKVKRLOSQRSPKSASELMSPSSRLMSLSTYTKWCKLLYCNGL	74
Db	21	LTALAGAAVEHCVMPEFDSVKTKMSGLCP-----	50
QY	75	EPLYLCNGARGCAIWFODPTRTGTMDAFAKYIYRHNSTRIMSGRLATLWVPATAIYE	134
Db	51	-----CPE-TRCPT-----PVHSLMSIVKREGLRPLRGVNVAAAGSMFAHLLTF	94
QY	135	TAYDOLAPFLCGR-ALTRSDIYAPVAVAGALARLGTVIISPLERTKLQAQHSYSELGA	193
Db	95	TYVERKMGVILTGNSAGHSNLTATYASGVATLTHDALMNAPEVVKQMGMAFSPYSSLE	154
QY	194	CVRTAIVAOGGNRSIMLGMPTALRDVPSALWYENELVKSWMGLRKPQDTSVGSFVA	253
Db	155	CARCIVYNRGVAAVYRSYTTQOLMANVPEQAIHMFSEYF---WQHLVLPBHKYDPKSHLIA	211
QY	254	GGISGTYAAVLTLPFDVYKTRQVQALGAMENVRNP-----LHDSFTWL-----LLRIR	303
Db	212	GGLAGGGLAAALITTPMDCKT-----VLTQOAAEADPNRRIRFLQARIYRIGSDAVATTIY	267
QY	304	AESGKGLFAGLFLPRITKAAPSCAIMISTYEFGK	337
Db	268	SORGLSGPSCGLQARIVTFQVYATATLMSVYELFK	301

```

RESULT 15
S54524
hypothetical protein YMR166c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein YM8520.15c
C:Species: Saccharomyces cerevisiae
C:Date: 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 29-Oct-1999
C:Accession: S54524; S54611
R:Hunt, S.; Bowman, S
submitted to the EMBL Data Library, May 1995
A:Reference number: S54510
A:Accession: S54524
A:Molecule type: DNA
A:Residues: 1-368 <HUN>
A:Cross-references: GB:249705; EMBL:249700; NID:9825556; PIDN:CAA9802.1; PTD:9825571
A:Experimental source: strain AB972
C:Genetics:
A:Map position: 13R
C:Superfamily: ADP/ATP carrier protein repeat homology
C:Keywords: duplication; transmembrane protein
F:501-139/Domain: ADP/ATP carrier protein repeat homology <ACP1>
F:145-241/Domain: ADP/ATP carrier protein repeat homology <ACP2>
F:253-361/Domain: ADP/ATP carrier protein repeat homology <ACP3>

```

Query Match	14.7%	Score 268.5	DB 2	Length 368	
Best Local Similarity	24.9%	Pred. No. 1.9e-15			
Matches	90	Conservative	56	Mismatches 136	Indels 79
				Gaps	9

QY	2	ADPDAGISFLDQOMAS	-CTGAVNVSLEPTPLDYYVRLQSORPSMASSELMPSRSLMSLS	60
		:-:-	:-:-	
		:-:-	:-:-	
DB	46	SDSD--LSPIMHCYVSGSIGIKIGDSAMHSLDYKTRDGG	-----	83
		:-:-	:-:-	
QY	61	YTKMKCLLYCNGVLEPLLYLCPNGARCATWFODPTRFTGMDAPVKIVRHEGFTT	-LMSGI	119
		:-:-	:-:-	
		:-:-	:-:-	
DB	84	-----APN-----	VKKRNNMISAVRTIMLEEGVRRGILYGY	114
		:-:-	:-:-	
QY	120	PATLVNTPAPATVETAYADQLKAFICGALTSDLVAPMAAGALARGVTVISPLEIMRT		179
		:-:-	:-:-	
		:-:-	:-:-	
DB	115	MAAMLGSPSALIFCTGYTYIKRTMIEDQINDTIHLSAGFLGRDIFSSFYVVPSEVLT		174
		:-:-	:-:-	
QY	180	KIQAQ-----	HVSRELGACVRYTAAVQGGMSLMLGMCPTLLRDPVPFSAALTWNY	229
		:-:-	:-:-	
		:-:-	:-:-	

Sat Apr 20 10:03:14 2002

us-09-840-787-19.rpr

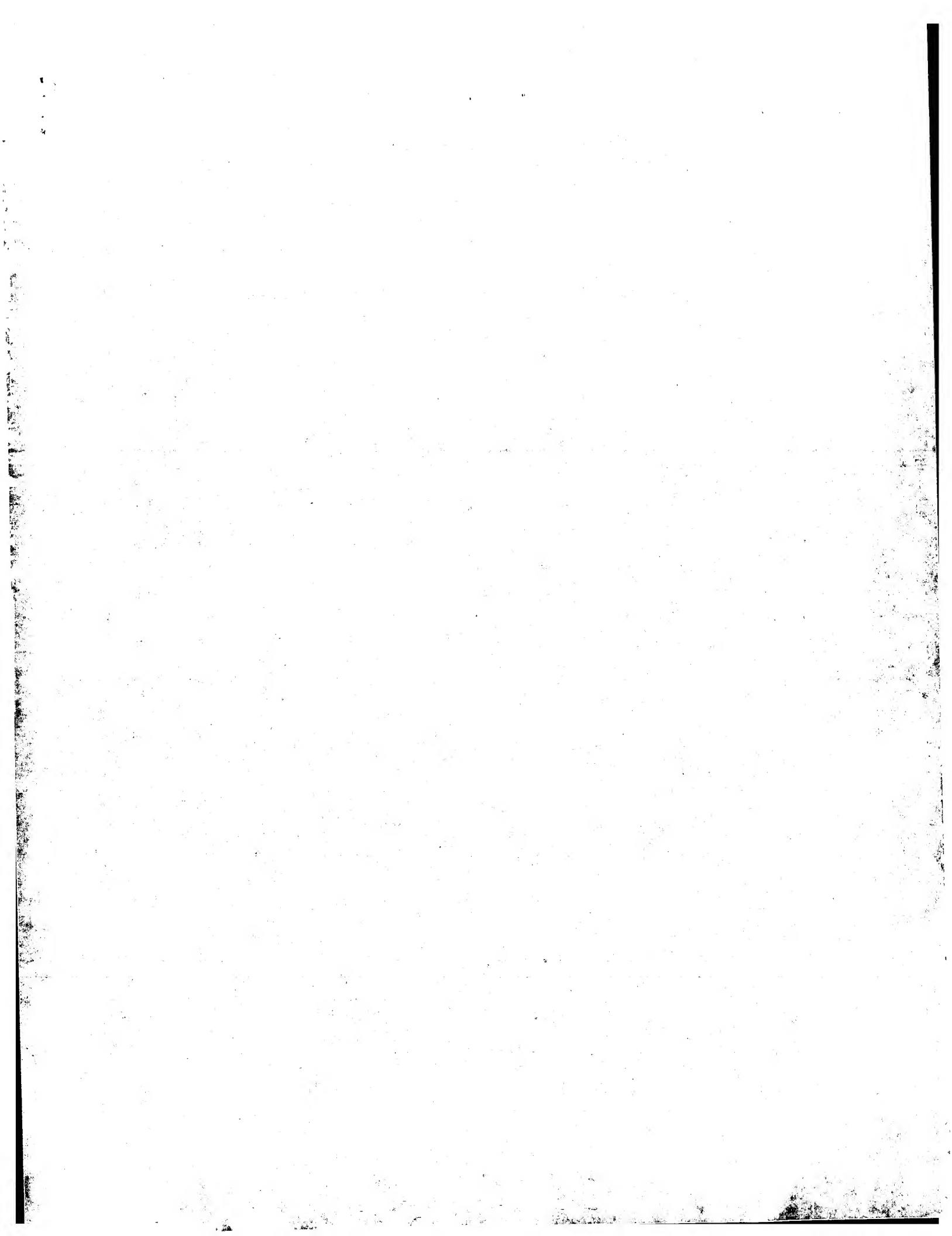
Page 7

```

Db      175 RQLOGRNNPFPOSGIWNYSULRNAITVINEEGFRSLFEGYKATLARDLPFSALGFAY 234
OY      230 ELVXSMWNLGRPKD---QTSVGMFSVAGGISGTVAAVLPLPDEVYKTO---QVALIGAM 282
        :      :      :      :      :      :      :      :      :      :
Db      235 EKFRPLAFRIQKQKGRDELSIPHEIITLGACAGGLAITTTTPMDVYKTRQTOQOPSOIN 294
OY      283 EAVRNPPLHV-----DSTWLLLRIRIAESGTKLFLFAGFLPRIIKAAPCAIMISY 333
        :      :      :      :      :      :      :      :      :      :
Db      295 KSYSTVTHPVHYNGRPALNSISLSLTVYQSEGVGLFESGVGPRFVWISVOSSILPLY 354
OY      334 E 334
        :
Db      355 Q 355

```

Search completed: April 19, 2002, 22:43:22
Job time: 3355 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2002, 22:42:02 ; Search time 68 Seconds

(without alignments)
189.255 Million cell updates/sec

Title: US-09-840-787-19

Perfect score: 1828

Sequence: 1 MADQDPAGISPLQOMVASGT.....TYEFGKSPFORLMDRLTGG 351

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Swissprot_39:*

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match Length	ID	Description
1	648.5	35.5	328 1	Y051_CAEEL Q09461 caenorhabdi
2	474.5	26.0	366 1	Y05F_YEAST P33320 saccharomyc
3	283	15.5	695 1	CMC1_DROME Q9ya73 drosophila
4	282.5	15.5	301 1	MCAT_RAT P97521 rattus norv
5	280	15.3	301 1	MCAT_HUMAN O43772 homo sapien
6	273	14.9	676 1	CMC2_MOUSE O9gx44 mus musculu
7	271	14.8	675 1	CMC2_HUMAN Q9aj50 homo sapien
8	268.5	14.7	368 1	Y039_YEAST Q03829 saccharomyc
9	264	14.4	702 1	CMC1_CAEEL Q21153 caenorhabdi
10	260.5	14.3	302 1	Y01K_SCHPO Q10248 schizosacch
11	259.5	14.2	307 1	Y07L_YEAST O99297 saccharomyc
12	255.5	14.0	310 1	UCP2_BRARE O94720 brachydanio
13	249	13.6	304 1	MRS4_YEAST P23500 saccharomyc
14	245.5	13.4	377 1	RIM2_YEAST P38127 saccharomyc
15	243.5	13.3	322 1	ACR1_YEAST P33303 saccharomyc
16	239	13.1	678 1	CMC1_HUMAN O73746 homo sapien
17	237	13.0	309 1	UCP2_HUMAN P55851 homo sapien
18	236.5	12.9	310 1	UCP2_CYPCA O94725 cyprinus ca
19	233	12.7	330 1	GDC_BOVIN Q01888 bos taurus
20	233	12.7	373 1	Y1A6_YEAST P40556 saccharomyc
21	232	12.7	309 1	UCP2_CANFA Q9a231 canis fami1
22	232	12.6	314 1	MRS3_YEAST P10566 saccharomyc
23	230	12.6	309 1	UCP2_RAT P56500 rattus norv
24	229	12.5	309 1	UCP2_MOUSE P70406 mus musculu
25	227.5	12.4	339 1	ADT_CHLKE P31692 chlorella k
26	226.5	12.4	284 1	PET8_YEAST P39921 saccharomyc
27	223	12.2	345 1	UCP3_RAT Q10442 schizosacch
28	221.5	12.1	308 1	YDEP_SCHPO P56499 rattus norv
29	221	12.1	436 1	B7L_MAIZE P29518 zea mays (m
30	220.5	12.0	323 1	UCP4_HUMAN O95847 homo sapien
31	219.5	12.0	308 1	UCP3_MOUSE P56501 mus musculu
32	219.5	12.0	357 1	YH62_YEAST P38702 saccharomyc
33	218	11.9	309 1	UCP2_PIG O97562 sus scrofa

34	213.5	11.7	292 1	AR11_YEAST Q12375 saccharomyc
35	212.5	11.6	313 1	ADT_NEUCR P02723 neurospora
36	210	11.5	307 1	UCP1_HUMAN P25874 homo sapien
37	209	11.4	288 1	UCP1_BOVIN P10861 bos taurus
38	206.5	11.3	311 1	UCP3_BOVIN O77792 bos taurus
39	205	11.2	318 1	ADT2_YEAST P18239 saccharomyc
40	205	11.2	332 1	GDC_HUMAN P16260 homo sapien
41	204	11.2	305 1	ADT_KLULA P49382 kluyveromyc
42	202.5	11.1	308 1	UCP3_PIG O97649 sus scrofa
43	201	11.0	311 1	UCP3_CANFA O9a239 canis fami1
44	199.5	10.9	298 1	ADT3_BOVIN P32007 bos taurus
45	199	10.9	335 1	YEA6_YEAST P39953 saccharomyc

ALIGNMENTS

RESULT	ID	Y051_CAEEL	STANDARD	PRT	328 AA.
AC	Q09461				
DT	01-NOV-1997	(Rel. 35, Created)			
DT	01-NOV-1997	(Rel. 35, Last sequence update)			
DT	01-NOV-1997	(Rel. 35, Last annotation update)			
DE	POTATIVE MITOCHONDRIAL CARRIER C16C10.1.				
GN	C16C10.1				
OS	Caenorhabditis elegans.				
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae; Rhabditidae;				
CC	Nematoda; Pelodermidae; Caenorhabditis.				
OK	NCBI_TaxID=6239;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=BRISTOL N2;				
RA	Lloyd C.;				
RL	Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.				
CC	-I- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.				
CC	-I- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
DR	EMBL: Z46787. CAAB6739.1. -				
DR	Wormpep; C16C10.1; CE01489.				
DR	Interpro; IPR001993; Mitoch_carrier.				
DR	Pfam; PF00153; mito_carr; 3.				
DR	PROSITE; PS00215; MITOCH_CARRIER; 1.				
KW	Hypothetical protein; Mitochondrion; Repeat; Transmembrane;				
KW	Transport.				
FT	TRANSMEM 93 113				POTENTIAL.
FT	TRANSMEM 146 166				POTENTIAL.
FT	TRANSMEM 237 257				POTENTIAL.
FT	SEQUENCE 328 AA; 36743 MW; EA1E9E329A764DE6 CRC64;				
Query Match	35.5%; Score 648.5; DB 1; Length 328;				
Best Local Similarity	39.4%; Pred. No. 2.7e-48;				
Matches 136; Conservative 63; Mismatches 113; Indels 33; Gaps 6;					
QY	10 SPLQOMVAGSGCAVVTSLFMPPLDVKYKRLDOO-PPSMASELMPSRLMSLTWKCLL 68				
DB	3 SPCESGKTTNCSGAPSSSCVPLDVKYKRLDOOQTRP-----FPKGEQFY 46				
QY	69 YCNGLEPLDYLCNGARATWFOPTRTGTMDAFVKIVRHGRTLTMSGLPATLVMTVP 128				
DB	47 YHNGLEHVCVCEVRKCEWYQRCNGRGTADLVKTAIRHGGISLWSGLSPTVMALP 106				
QY	129 ATATFTAYDQAKFL-----CGRALTSDLTAP-----WAGALARIGTVIVIPLELM 177				
DB	107 ATVFYFTTYDNLVWLKKRMCRCRAFSPEKWTVPDMSAAVAGIVARTIATVVSPIEMI 166				

RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mekulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Paolel J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard D.R., Paoletti J.M.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svayras R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
CC -1- FUNCTION: CALCIUM-DEPENDENT MITOCHONDRIAL SOLUTE CARRIER (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC INNER MEMBRANE (BY SIMILARITY).
CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS: 1 (SHOWN HERE), 2 AND 3; MAY BE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY. ARALAR
CC SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 4 EF-HAND CALCIUM-BINDING DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Y18197. CAB62169.1; -;
DR EMBL: AE003774; AAF57048.1; -;
DR EMBL: AE003774; AAF57049.1; -;
DR EMBL: AE003774; AAF57050.2; -;
DR -FlyBase: FBgn028646; atalari.
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR002067; Mit_carrier.
DR InterPro: IPR001993; Mitoch_carrier.
DR Pfam: PF00036; efhand; 3.
DR Pfam: PF00153; mito_carr; 3.
DR PRINTS: PR00926; MITOCARRIER.
DR SMART: SM00054; EFh; 3.
DR PROSITE: PS00018; EF-HAND; 1.
DR PROSITE: PS00215; MITOCH_CARRIER; 3.
KW Calcium-binding; Alternating; Repeat; Transmembrane; Transport;
KW Calcium-binding; Alternating; Repeat; Transmembrane; Transport;
FT TRANSMEM 346 425 1 (POTENTIAL).
FT TRANSMEM 406 425 2 (POTENTIAL).
FT TRANSMEM 449 462 3 (POTENTIAL).
FT TRANSMEM 498 517 4 (POTENTIAL).
FT TRANSMEM 537 554 5 (POTENTIAL).
FT TRANSMEM 594 613 6 (POTENTIAL).
FT CA_BIND 84 95 EF-HAND 1.
FT CA_BIND 118 129 EF-HAND 2.
FT DOMAIN 157 168 ANCESTRAL CALCIUM SITE 3.
FT CA_BIND 189 200 EF-HAND 4.
FT DOMAIN 225 235 3 APPROXIMATE TANDEM REPEATS.
FT REPEAT 225 434 1.
FT REPEAT 435 524 2.
FT REPEAT 525 627 3.
FT VASPLIC 1 52 MISSING (IN ISOFORM 3). LTRSLPN (IN
FT VASPLIC 3 22 ISOFORM 2).
FT GA -> AP (IN REF. 1).
FT CONFLICT 356 357 V -> A (IN REF. 1).
FT CONFLICT 556 556 V -> L (IN REF. 1).
FT CONFLICT 653 653 V -> L (IN REF. 1).
SQ SEQUENCE 695 AA; 76753 MW; 8EC93D92031F5B77 CMC64;

Query Match 15.5%; Score 283; DB 1; Length 695;
Best Local Similarity 27.3%; Pred. No. 1,2e-16;
Matches 97; Conservative 55; Mismatches 129; Indels 74; Gaps 12;
OY 2 ADQDPAGISPLQOMVAGS-----TGAVTSLEMTPLDVVKVRLQSGRPSMASLMS 53
DB 326 AVEPSADRSFAQVLESSRYFTLSFAGAVGATV--VYPIDLVKTRMQNR-----AG 376
OY 54 SLMNLSY-TRKCLLYCNGVLEPLICNGARCATWEDPFRFTGTMDAPKRYHSET 112
DB 377 SYIGEVAVRNSWDC-----FRKVVYRHEGF 400
OY 113 RFLMSGPLPATIYFWPATIYFTAYDQKAFICGRALTSDLAPVAGALARGTYVIS 172
DB 401 MCLYKGLPLQMGVAPERAKITVDLVNDKDKGNIPYAEVAGGACASQVFTN 460
OY 173 PLEMRKTLQ-AQHSYRELACVR--TAVAGGMRSLWLGMPALRDVPSALWPNY 229
DB 461 PLEIKIRIQVAGEIA---SGSKIRAMSVRELGLFGYKGRACLLRDVPSALYFTY 517
OY 230 ELVSWLNGLRKQDOTSGMSFY-AGISGTYAAVLTLPDVKYQROVALCAMEAVRN 288
DB 518 AHTKAM---MADKQVNHPLTLAAGALAGVPAASIVPADVIKTRLDV-----VARSG 568
OY 289 PLHVDSTMLLRIRAESGTGLFAGFLPRIKAPSCAIMSTYEFKSFQRL 343
DB 569 QTTTGVWDATKKIMAESEPRAFMKGTARVRSRPSQGVLTYTE---LLQRL 619
RESULT 4
MCAT_RAT
AC P97521. STANDARD. PRT: 301 AA.
DI 01-NOV-1997 (Rel. 35, Created)
DI 01-NOV-1997 (Rel. 35, Last sequence update)
DI 30-MAY-2000 (Rel. 39, Last annotation update)
DE MITOCHONDRIAL CARNITINE/ACYLCARNITINE CARRIER PROTEIN
DE (CARNITINE/ACYLCARNITINE TRANSLOCASE) (CAC).
GN CACT.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RP [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=97184657; PubMed=9032458;
RA Indiveri C., Jacobazzi V., Giangregorio N., Palmieri F.;
RT "The mitochondrial carnitine carrier protein: cDNA cloning, primary
RT structure and comparison with other mitochondrial transport
RT proteins.";
RL Biochem. J. 321:713-719(1997).
CC -1- FUNCTION: TRANSPORT OF CARNITINE OR ACYLCARNITINE FROM THE CYTOSOL
CC TO THE MITOCHONDRIAL MATRIX.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC INNER MEMBRANE.
CC -1- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -1- PTR: THE N-TERMINAL IS BLOCKED.
CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X97831; CAA66410.1; -;
DR InterPro: IPR001993; Mitoch_carrier.
DR Pfam: PF00153; mito_carr; 3.
DR PROSITE: PS00215; MITOCH_CARRIER; 3.

Query Match	Best Local Similarity	Score	DB 1:	Length	301:
Matches	93: Conservative	52: Mismatches	130: Indels	67: Gaps	
QY 4	ODPAPISPLQOOWVAGSTGAVVNTSLMPTPLADYKVRLOSORPMSASELMPSSRLMSLSYRK	63			
DB 3	EEPRKPSLEPKNLNLAGFGGVCVLPVGHPLDYKVKALQTPSPSLPQC	48			
QY 64	WKCILLYCNGVLEPLLYLCPNGARCAWTFODPTREFTGTMDAIVKIVHNEGTRTLMGIPATL	123			
DB 49	-----PRMYSGLTIDCFRKLTRERGTITGLYRGMAPI	79			
QY 124	WVTVAATATVETPANDOLKALFGLRALTSLDPAWV--AGALRLGTVWYISPLEMRTKL	181			
DB 80	IGVTMPFAVCFEGFG-LGKRLQOKSPDEDELITYPOLFTBMLSGVTTGTMTGERIKCLL	138			
QY 182	QAQHVY---YRELGACVRYTAAVAGGWRSLMLGWPALRDVPFSAIWNELVKSMLN	237			
DB 139	QIQASSGKNKKYSGLTDCAKIKLYQERIGFYKGTALTLMRDVPASOMYEMTE---WLK	194			
QY 238	GL-RRKDOT---SYGMSVYAGGISGTYAAVLTLPEDYVKTRQVVALGMEAVRNPPLHV	292			
DB 195	NLETFQGSVHDLSPVRLVYVAGGFGIENWVAIPDVLKSRFQAAPPK-----YP	246			
QY 293	DSTWLLRLRRIRASGCTGLFAGFLPRRIKAAPSCAIMISTYE	334			
DB 247	NGFRDVLRELIREEGVTSLYKGFNAVMIRAFANNAACFLGFE	288			

```

CC CC INNER MEMBRANE.
CC -1- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -1- DISEASE: DEFECTS IN CACT ARE THE CAUSE OF AN AUTOSOMAL RECESSIVE
CC DEFICIENCY IN MITOCHONDRIAL OXIDATION OF FATY ACIDS.
CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
CC DR EMBL, Y10319; CAA71367.1; -.
CC DR MIM, 212138; -.
CC DR InterPro: IPR001993; Mitoch_carrler.
CC DR Pfam, PF00153; MITOCH_CARRIER; 3.
CC DR PROSITE, PS00215; MITOCH_CARRIER; 3.
CC FT Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport.
CC KM DOMAIN 1 12 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 13 31 POTENTIAL.
CC FT DOMAIN 32 73 MITOCHONDRIAL MATRIX (POTENTIAL).
CC FT TRANSMEM 74 93 POTENTIAL.
CC FT TRANSMEM 94 112 CYTOPLASMIC (POTENTIAL).
CC FT DOMAIN 113 131 POTENTIAL.
CC FT TRANSMEM 132 170 MITOCHONDRIAL MATRIX (POTENTIAL).
CC FT TRANSMEM 171 190 POTENTIAL.
CC FT DOMAIN 191 211 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 212 230 POTENTIAL.
CC FT DOMAIN 231 267 MITOCHONDRIAL MATRIX (POTENTIAL).
CC FT TRANSMEM 268 287 POTENTIAL.
CC FT DOMAIN 288 301 CYTOPLASMIC (POTENTIAL).
CC FT SEQUENCE 301 AA; 32943 MW; AEB34E4E35102B0 CRC64;
CC SQ

```

DE FAMILY 25, MEMBER 13) (CITRIN).
 CC SLIC25A13 OR ARALAR2.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RX MEDLINE=20079163; PubMed=10610724;
 RA Sinasac D.S., Crackower M.A., Lee J.R., Kobayashi K., Saheki T.,
 RA Scherer S.W., Tsai L.-C.;
 RT "Genomic structure of the adult-onset type II citrullinemia gene,
 RT SLIC25A13, and cloning and expression of its mouse homologue.";
 RL Genomics 62:289-292(1999).
 CC -1- FUNCTION: CALCIUM-DEPENDENT MITOCHONDRIAL SOLUTE CARRIER. MAY HAVE
 CC A FUNCTION IN THE UREA CYCLE (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
 CC INNER MEMBRANE (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: AT E10.45, EXPRESSED IN BRANCHIAL ARCHES AND
 CC LARIMB AND TAIL BUDS. AT E13.5 EXPRESSION IS PREDOMINANT IN
 CC EPITHELIAL STRUCTURES AND THE FOREBRAIN, KIDNEY AND LIVER.
 CC EXPRESSION IN LIVER IS MAINTAINED INTO ADULTHOOD.
 CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY. ARALAR
 CC SUBFAMILY.
 CC -1- SIMILARITY: CONTRAINS 4 EF-HAND CALCIUM-BINDING DOMAINS.
 CC -----
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@isb-sdb.ch).
 CC -----
 CC DR EMBL, AF164633; AAF21426.1; -
 CC DR MGD; MGI:1354721; SLIC25a13.
 DR InterPro: IPR002048; EF-hand.
 DR InterPro: IPR002067; Mito_carrier.
 DR InterPro: IPR001993; Mitochn_carrier.
 DR -pfam; PF00036; ethand; 3.
 DR -pfam; PF00153; mito_carr; 3.
 DR PRINTS: PRO0926; MITOCARRIER.
 DR PROSITE: PS00018; EF_HAND; FALSE_NEG.
 DR PROSITE: PS00215; MITOCH_CARRIER; 3.
 KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport;
 KW Calcium-binding.
 FT TRANSMEM 333 350 1 (POTENTIAL).
 FT TRANSMEM 394 413 2 (POTENTIAL).
 FT TRANSMEM 437 450 3 (POTENTIAL).
 FT TRANSMEM 486 505 4 (POTENTIAL).
 FT TRANSMEM 525 542 5 (POTENTIAL).
 FT TRANSMEM 582 601 6 (POTENTIAL).
 FT CA_BIND 66 77 EF-HAND 1.
 FT CA_BIND 100 111 EF-HAND 2.
 FT DOMAIN 139 150 ANCESTRAL CALCIUM SITE 3.
 FT CA_BIND 171 182 EF-HAND 4.
 FT DOMAIN 207 432 3 APPROXIMATE TANDEM REPEATS.
 FT REPEAT 207 432 1.
 FT REPEAT 423 512 2.
 FT REPEAT 513 615 3.
 SQ SEQUENCE 676 AA: 74466 MW: 5E395208B5E58C5 CRC64:

	Query Match	14.98;	Score 273;	DB 1;	Length 676;	
	Best Local Similarity	27.48;	Pred. No. 8.3e-16;			
	Matches 90;	Conservative	48;	Mismatches 130;	Indels 60;	Gaps 7.
QY	16 VASGTGAVVTSLEMTPLDVKVRLOSORP--SWASELMPSSRLMSLSYTKMKCLLYCNGV	73				
	: : :: : : : : : : :					
Dd	338 ICAAVNGA--TAAV--PDIDLKTRMQRKSTSGSVGLM-----	371				
QY	74 LEELPLYLCIPNARGCATWFODPRTFTGTMDIAFVKIVRIHGGTFLTWSGLPATLVMTVPATAIY	133				
	::::: :					

```

Db 372 - - - - - YKNSFDCKFKVRLRYEEFFCLYGLSLRQJLQVGAPEKAIK 409
Oy 134 FTAYDQKAFELCRALTSDLAPVWAGALARLGLVTYVISPLEMRKYLQAHVSYRELGA 193
Db 410 LTVNDEFRDCEFMKDKDSVPLLEIEIFGAGCGAGSGQVIFNPPEIKIRLQVAGEITTTGPV 469
Oy 194 CVRFAVAGGWRSLNLGNGPTRLRDVPPSALWFWNYELVK-SWTLNGLRPRPDOTSVGMSFV 252
Db 470 SALSIVVDLDFECITYGAKACFRLDRDIPSAITPCYIAHVKSFPN---EDGQVSPGSL 526
Oy 253 AGCISGTVAAVLTLPPDVYQTOHQVALGMAEVRVNPPLHVDSTMILLRIRAESGTGILE 312
Db 527 AGALTAGMPASLIVTPADVITKTRLDV-----AARAGQTYNGVYDCRKLIREGPKALM 580
Oy 313 AGFLPRITKAAPSCAIMISTYERGSKFE 340
Db 581 KGVARVFRSSPQFGVTLTYELLQWFE 608

RESULT 7
CMC2_HUMAN STANDARD: PRT: 675 AA.
ID CMC2_HUMAN Q9UJS0: Q9UJN17; Q9VZM1.
AC 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CALCIN-BINDING MITOCHONDRIAL CARRIER PROTEIN ARLAR2 (SOLUTE CARRIER
DE FAMILY 25, MEMBER 13) (CITRIN).
GN SLC25A13 OR ARLAR2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid:9606;
RN [1]
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RX MEDLINE=99259531; PubMed=10369257;
RA Kobayashi K., Sinasac D.S., Iijima M., Boright A.P., Begum L.,
RA Lee J.R., Yasuda T., Ikeda S., Hiltano R., Terrazono H., Crackower M.A.,
RA Kondo I., Tsui L.-C., Scherer S.W., Sahelki T.;
RA "The gene mutated in adult-onset type II citrullinaemia encodes a
RT putative mitochondrial carrier protein.";
RL Nat. Genet. 22:159-163(1999).
RN [2]
RN SEQUENCE FROM N.A., AND CHARACTERIZATION.
RP TISSUE=Liver;
RC MEDLINE=20115110; PubMed=10642534;
RA Del Arco A., Aguado M., Satriestegui J.;
RA "Characterization of a second member of the subfamily of calcium-
RT binding mitochondrial carriers expressed in human non-excitabile
RT tissues.";
RL Biochem. J. 345:725-732(2000).
RN [3]
RN SEQUENCE OF 6-251 FROM N.A., AND DISEASE.
RX MEDLINE=20079163; PubMed=10610724;
RA Sinasac D.S., Crackower M.A., Lee J.R., Kobayashi K., Sahelki T.,
RA Scherer S.W., Tsui L.C.;
RA "Genomic structure of the adult-onset type II citrullinaemia gene,
RT SLC25A13, and cloning and expression of its mouse homologue.";
RL Genomics 62:289-292(1999).
CC -1- FUNCTION: CALCIUM-DEPENDENT MITOCHONDRIAL Solute CARRIER. MAY HAVE
CC A FUNCTION IN THE UREA CYCLE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC INNER MEMBRANE.
CC -1- TISSUE SPECIFICITY: HIGH LEVELS IN LIVER AND LOW LEVELS IN KIDNEY,
CC PANCREAS, PLACENTA, HEART AND BRAIN.
CC -1- DISEASES: DEFECTS IN SLC25A13 ARE A CAUSE OF ADULT-ONSET TYPE-II
CC CITRULLINAEMIA (CTLN2). CTLN2 IS CHARACTERIZED BY A LIVER SPECIFIC
CC DEFECT IN THE REGULATION OF ARGININOSUCCINATE SYNTHASE (ASS)
CC RESULTING IN A DEFICIENCY OF ASS. A FEW YEARS AFTER ONSET PATIENTS
CC ENTER A COMA AND SUFFER CEREBRAL EDEMA.
CC -1- MISCELLANEOUS: BINDS CALCIUM.
CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY. ARLAR
CC SUBFAMILY.

```


Db 235 EKEFOLAEKIEQKGRDGLSLPNEILTGACAGGLAGITTPMDVYKTRVOTQPPSOSN 294

QY 283 EAVRNPLHV-----DSTWLLLRIRAESGTGKLFAGFLPRIRKAPSCAIMISTY 333

Db 295 KSYSTHPRHVTRNGRPAALNSISLSLRVYQSEGLVGFSSGVPFRVWTSVSSIMLLY 354

QY 334 E 334

Db 355 Q 355

RESULT 9

CMC1_CAEEL STANDARD: PRT: 702 AA.

AC Q21153; 20-AUG-2001 (Rel. 40, Created)

DT 20-AUG-2001 (Rel. 40, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE PROBABLE CALCIUM-BINDING MITOCHONDRIAL CARRIER K02F3.2.

GN K02F3.2.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

CC Rhabditidae; Peloderinae; Caenorhabditis.

CC NCB1_TaxID=6239;

RA NCBI_TaxID=6239;

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RA Bentley D.;

CC Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: CALCIUM-DEPENDENT MITOCHONDRIAL SOLUTE CARRIER (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL INNER MEMBRANE (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY. ARAALAR SUBFAMILY.

CC -1- SIMILARITY: CONTAINS 4 EF-HAND CALCIUM-BINDING DOMAINS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: U00052; AKK21421.1; -

DR Wotmpep; K02F3.2; GE01348.

DR InterPro: IPR002048; EF-hand.

DR InterPro: IPR001993; Mitoch_carrler.

DR Pfam: PF00036; efhand; 3.

DR Pfam: PF00153; mito_carr; 3.

DR PRINTS: PR00926; MITOCARRIER.

DR PRINTS: PR00927; ADPTRNSLCASE.

DR SMART: SM00054; EFh; 3.

DR PROSITE: PS00018; EF_HAND; 2.

DR PROSITE: PS00215; MITOCH_CARRIER; 3.

KW Hypothetical protein: Mitochondrion; Inner membrane; Repeat.

KM Transmembrane; Transp; Calcium-binding.

FT TRANSMEM 368 385

FT TRANSMEM 429 448

FT TRANSMEM 471 484

FT TRANSMEM 521 540

FT TRANSMEM 560 577

FT TRANSMEM 617 636

FT CA_BIND 106 117

FT CA_BIND 140 151

FT CA_BIND 179 190

FT DOMAIN 211 222

FT DOMAIN 247 252

FT REPEAT 247 252

FT REPEAT 457 457

FT REPEAT 548 548

FT REPEAT 650 650

FT REPEAT 702 702

SEQUENCE 702 AA; 78170 MW; 41AB9EF9A1F018D0 CRC64;

Query Match. 14.4%; Score 264; DB 1; Length 702;

Best Local Similarity 27.3%; Pred. No. 5.1e-15;

Matches 91; Conservative 46; Mismatches 128; Indels 68; Gaps 9;

QY 16 VASGTAVVTSLEMTPLDVVKVRLQSORPSMASSELPSSRLSTSYTKWICLLYGVLE 75

Db 373 VAGACGA--TANY--PIDLVKTRMGNKRTSS----- 400

QY 76 PLYLCPNGARCATWFDPPRTGTMDAIVKIVRHGRTLTMSGLPATLVMTATATYFT 135

Db 401 -----FVGEVMYKNSIDCEKVKYKFEGLLIGRLPQIVGVAPEKAIKLT 446

QY 136 AVDOAKFLCGRALSD----LYAPMVGALARLSTVIVISLELMRKLO AQHVSIRE 190

Db 447 MNDYMR-----DKFKDKKIPLYGELIAGTGMCOVVFNPTELKIRLQTAGVQOAG 501

QY 191 LGACVRTVAOQGMRSMLMGWPTLRDVPFSALYFNVYELKSWLNGLRPKDQTSVGM 250

Db 502 KKIGFTYVLKELGFLGLYKSGRACFLRDIPIFAIYPAVNAHK--LASADEGMSPTL 559

QY 251 FVAGISGTVAIVTLPPDVYKTRQVALGANEAVRNPLHVDSTWLLLRIRAESGTGK 310

Db 560 FASAFIAGVPAAGLVTPADVITKRLQV-----AARAGQTYNGVIDCARLKIREGPM 613

QY 311 LFAGFLPRIRKAPSCAIMISTYERKSEFQRL 343

Db 614 LMKGTAAVCRSSPQFAVTLTYE---VLQRL 642

RESULT 10

YDK_SCHPO STANDARD: PRT: 302 AA.

AC Q10248;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 01-OCT-1996 (Rel. 34, Last annotation update)

DE PURNATIVE MITOCHONDRIAL CARRIER C4G9.20C.

GN SPAC4G9.20C.

OS Schizosaccharomyces pombe (Fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

CC Schizosaccharomycetes; Schizosaccharomycetaceae;

CC NCB1_TaxID=4896;

RA NCBI_TaxID=4896;

RP SEQUENCE FROM N.A.

RC STRAIN=972;

RA Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;

RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL INNER MEMBRANE (POTENTIAL).

CC -1- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.

CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: Z69727; CA93570.1; -

DR InterPro: IPR001993; Mitoch_carrler.

DR Pfam: PF00153; mito_carr; 3.

DR PROSITE: PS00215; MITOCH_CARRIER; 2.

KW Hypothetical protein: Mitochondrion; Inner membrane; Repeat.

KM Transmembrane; Transp.

FT TRANSMEM 9 29

FT TRANSMEM 71 91

FT TRANSMEM 116 136

FT TRANSMEM 216 236

SO	SEQUENCE	302 AA;	33348 MM;	FF7EB8A542E61C481 CRC64;
Query Match	14.3%;	Score 260.5;	DB 1;	Length 302;
Best Local Similarity	25.5%;	Pred. No. 3.8e-15;		
Matches 84;	Conservative 53;	Mismatches 130;	Indels 63;	Gaps 8;
QY	13	QOMVASGCAVVTSLFMTPLDVVKYRLOSQRPSNASLSMPSRLMSLSYTKMKCLLYONG	72	
Db	19	KDFLAVSGGVAQYVLVGPEPCVKRLQSQ-----S	49	
QY	73	VLEPLVLYCPNRCARCATWQDPTRFETGTDAAVKIVRHGTRTLMSSGLPATLMTVPATAI	132	
Db	50	NVSPILY-----NNALDVCRKISKSNGLAFAFYGVLPILLGIFCVSI	91	
QY	133	YFTAYDQLKALC--GRALTSDDLAPMVAAGALARIQYTVISPLELMTKIQAO---HVS	187	
Db	92	QFTFEYCKRFRSSRDGTFTVMPQY--YVSGAISGLANSPLVGPRVHVRIRLDIQGNVL	149	
QY	188	YRELGACVRYTVAOAGGMSLIMGCPTRLARDVPFSALTFWFEY--LVKSML--NGLRPKDI	244	
Db	150	YHGPDICLKRTSSQYGLSGIMKGYPTLAREHNGMFLFYEALVYKNTMAKHHLLTDSQ	209	
QY	245	TSVGSFVAAGISGTVAVALTLPEDVYKTRQKQVALGAEAVRNPDLHDSITWLLIRIRA	304	
Db	210	TPGWLKCFYCGAGAGAMMLAAYPPPIVYSKIOTDGLSKATYKN-----SMQAKGIYT	263	
QY	305	ESGTFGLFAGFLPRITKAAPSCAIMISITYE	334	
Db	264	KAGLGFYRGFVPLVYRAAPANAATFYIYE	293	
RESULT 11				
ID	Y07L_YEAST	STANDARD;	PRT;	307 AA.
AC	Q99297;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DE	POTATIVE MITOCHONDRIAL CARRIER YOR222W.			
GN	YOR222W OR YOR50-12.			
OS	Saccharomyces cerevisiae (Baker's yeast).			
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;			
CC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.			
CC	NCBI_TaxID=4932;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=S288C / FY1679;			
RC	MEDLINE=96437977; PubMed=8840505;			
RA	Gallison F., Dujon B.;			
RT	"Sequence and analysis of a 33 kb fragment from the right arm of			
RT	chromosome XV of the yeast Saccharomyces cerevisiae."			
RL	Yeast 12:877-885(1996).			
CC	-1 SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.			
CC	-1 MITOCHONDRIAL INNER MEMBRANE (BY SIMILARITY).			
CC	-1 DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.			
CC	-1 SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@sib-sib.ch).			
CC	-----			
DR	EMBL; Z75130; CAA99440.1; -			
DR	EMBL; X92441; CAA63185.1; -			
DR	SGD; S0005748; YOR222W.			
DR	InterPro; IPR002067; Mit_carrier.			
DR	InterPro; IPR001993; Mitoch_carrier.			
DR	Pfam; PF00153; mito_carr; 3.			
DR	PRINTS; PR00926; MITOCARRIER.			

```

DR PRINTS:PPR00927; ADPTRNSLCASE.
DR PROSITE, PS00215; MITOCH_CARRIER_3.
KW Hypothetical protein, Mitochondrion; Inner membrane; Repeat;
KW Transmembrane; Transpos.
FT TRANSMEM 10 POTENTIAL.
FT TRANSMEM 122 142 POTENTIAL.
FT TRANSMEM 215 235 POTENTIAL.
FT TRANSMEM 280 300 POTENTIAL.
SQ SEQUENCE 307 AA; 34007 MW; 4089082A64DBA97C CRC64;

Query Match 14.2%; Score 259.5; DB 1; Length 307;
Best Local Similarity 24.6%; Pred. No. 4,8e-15;
Matches 85; Conservative 56; Mismatches 156; Indels 49; Gaps 6;

QY 2 ADPDAGISPLDQWASGSGAVVSLFEMTPIDVYVVRLOSQRPMSASLMSRLMSLSY 61
DB 3 SDSNKKPLEFTYQFISGAVAGISLFTVMPIDVYVTRQLE----- 43
QY 62 TWKCLLYCNGVLEPLDYLCPNGARCATFODPRETGTMDAFVKIVRHEGRTLWSGLPA 121
DB 44 -----VTP-----TAAVGQKQERYNGVIDCLAKIVKKEGFSRLRGIS 84
QY 122 TLVMTVPATATFTYADOLKAFLCGRALTSGLD--YAPVAVAGALATGVYISPLEMRT 179
DB 85 PMLMEAPKRAFTEACNDQYOKTEKNLFENNETOKISIAAGSASMTAAAYVPELTKI 144
QY 180 KIQAOHVSYRELACVRTAVAGQWRSIMLGWGPALRDVPSALYWNELVYSWMLNG 239
DB 145 RNQDVKSSYLPGMDCLKIKINEGLMGLEYKGIESTFMERNALNGSGYFGVITYQVRNSMPYA 204
QY 240 RKRDTSVMSYVAGISGTVAAYLTLPDYVYKTRQYVALGAMENVRNPLHYDSTWL- 298
DB 205 KTKGQKTRN-DLIAAGIGVTGTMTPDYVKSISQSDAVASSVK-----KYNMCLP 257
QY 299 -TRRIAESGTGKLFAGLEPLRIKAPSCAAMISYEEGKSPFFQRL 343
DB 258 SILVYIREGPRALVKGFPVKYCRALPAGSGLMLVFTGMNMFFRDL 303

UCP2_12
UCP2_BRARE
ID UCP2_BRARE STANDARD; PRT; 310 AA.
AC 09AW20;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE MITOCHONDRIAL UNCOUPLING PROTEIN 2 (UCP 2).
GN UCP2.
OS Brachydanio rerio (zebrafish) (zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Rasbora; Danio.
OX NCBI_TaxId=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Stuart J.A., Harper J.A., Brindle K.M., Brand M.D.;
RT "Uncoupling protein 1 homologues and the thermogenesis of cold-
RT blooded vertebrates.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBD databases.
CC -!- FUNCTION: UCP ARE MITOCHONDRIAL TRANSPORTER PROTEINS THAT CREATE
CC PROTON LEAKS ACROSS THE INNER MITOCHONDRIAL MEMBRANE. THUS
CC UNCOUPLING OXIDATIVE PHOSPHORYLATION FROM ATP SYNTHESIS. AS A
CC RESULT, ENERGY IS DISSIPATED IN THE FORM OF HEAT (BY SIMILARITY).
CC -!- SUBUNIT: ACTS AS A DIMER FORMING A PROTON CHANNEL (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC INNER MEMBRANE.
CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation at
CC the European Bioinformatics Institute. There are no restrictions on its

```

"MRS3 and MRS4, two suppressors of mtRNA splicing defects in yeast, are new members of the mitochondrial carrier family."

```

0b      264 LEVHGWKGEWRT.KPRIVANIPATATISWTAVECAKUEI WY 202
      | | : | | : | | : | | :

```

264 LEVHGKGEWBGTKBPTVANTPATATSWTAVECAVUEHMY 303

```

RESULT 14
RIM2_YEAST STANDARD; PRT; 377 AA.
ID RIM2_YEAST
AC P38127;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE MITOCHONDRIAL CARRIER PROTEIN RIM2.
GN RIM2 OR YBR192W OR YBR140Z.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=93348777; PubMed=8346681;
RA Demolis N., Mallet L., Bussereau F., Jacquet M.;
RT "RIM2, MS1 and PGI1 are located within an 8 kb segment of
RT Saccharomyces cerevisiae chromosome II, which also contains the
RT putative ribosomal gene L21 and a new putative essential gene with a
RT leucine zipper motif."
RT Yeast 9:645-659(1993).
RL
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z21487; CAA9678.1; -
DR EMBL: Z36061; CAA85154.1; -
DR PIR: S36081; S36081.
DR SGD: S0000396; RIM2.
DR InterPro: IPR001993; Mitoch_carrier.
DR Pfam: PF00153; mito_carr_3.
DR PROSITE: PS00215; MITOCH_CARRIER; 1.
DR Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport.
KW SEQUENCE 377 AA; 42101 MW; A3C61383C3B3743B CRC64;
SQ
Query Match 13.4%; Score 245.5; DB 1; Length 377;
Best Local Similarity 22.2%; Pred. No..9.5e-14; Indels 109; Gaps 11;
Matches 84; Conservative 58; Mismatches 127;

```

```

OY 317 PRIKAPSCAISTEYE 334
ID PRIKAPSCAISTEYE
AC P33303;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE REGULATOR OF ACETYL-COA SYNTHETASE ACTIVITY.
GN ACRI OR YJR095W OR J1921.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN (1)
RP SEQUENCE FROM N.A.
RC MEDLINE=94203187; PubMed=7908717;
RX Fernandez M., Fernandez E., Rodicio R.;
RA "ACRI, a gene encoding a protein related to mitochondrial carriers,
RT is essential for acetyl-CoA synthetase activity in Saccharomyces
RT cerevisiae."
RT Mol. Gen. Genet. 242:727-735(1994).
RL
CC -1- INDUCTION: BY ETHANOL OR ACETATE AS SOLE CARBON SOURCES. REPRESSED
CC BY GLUCOSE.
CC -1- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z25485; CAA80973.1; -
DR EMBL: Z49595; CAA89624.1; -
DR PIR: S36407; S36407.
DR PIR: S43280; S43280.
DR SGD: S0003856; SFC1.
DR InterPro: IPR001993; Mitoch_carrier.
DR Pfam: PF00153; mito_carr_3.
DR PROSITE: PS00215; MITOCH_CARRIER; 3.
DR Transmembrane; Transport; Repeat.
KW CONFLICT 300 322
FT MTHOSOSH (IN REF. 1)
SQ SEQUENCE 322 AA; 35340 MW; EC29718A0F5011A7 CRC64;
Query Match 13.3%; Score 243.5; DB 1; Length 322;
Best Local Similarity 22.1%; Pred. No..1.2e-13;
Matches 81; Conservative 55; Mismatches 129; Indels 101; Gaps 9;

```

```
QY 127 VPATAIYFTAYDQLKAPL-----CGRALTSDLVAPVAVAGALARLGTVTYISPLEIMRTXLO 182
Db 83 IPKMAIRFSSYEFYRLLVKNKESGYSTGNTFYAGVAGITE--AVLVVPMFEVVKIRLO 140
QY 183 AQHVSYRELGA-----CVRTAVAQGGWMSLMLGMPALRDVPFSALYWFNYELVK 233
Db 141 AQHLPPEPNAGPKYNNAIHAAVTYKKEGVSALYRGVSLTAARQATNOGANFTYYSKLK 200
QY 234 SMLNGLRPKDQTSVGMSTFVAGISGTVAVALTLPEDVVKTORQ-----V 277
Db 201 EFLQNTHOMDVLPBSETSCIGLISGAIQPPSNAPLDTIKTRLOKDKSISLEKOSGMKKII 260
QY 278 ALGAMBAVRVNPPLHVDSTWLLLRIRARESGTKGLFAGFLPRIIKAPSCAIMISTYEFGK 337
Db 261 TIGA-----QLKEGFRALYKGTTPRVRYAPGQAVTFTYIEYVR 301
QY 338 SFFQRL 343
Db 302 EHLENL 307
```

Search completed: April 19, 2002, 22:52:36
Job time: 634 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2002, 21:46:02 ; Search time 62.21 Seconds
(without alignments)
126.968 Million cell updates/sec

Title: US-09-840-787-19

Perfect score: 1828

Sequence: 1 MADQDPAGISPLQOMVASGT.....TYEFGKSFQRLNDRLLG 351

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 segs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA: *
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep: *
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep: *
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep: *
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep: *
5: /cgn2_6/ptodata/2/1aa/PTOUS_COMB.pep: *
6: /cgn2_6/ptodata/2/1aa/Backfilest.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1828	100.0	351	2	US-08-933-750C-19
2	1828	100.0	351	4	US-09-234-613-19
3	324	17.7	328	4	US-09-068-140A-15
4	320	17.5	289	1	US-09-068-140A-10
5	237	13.0	299	1	US-08-518-878B-56
6	237	13.0	299	2	US-08-470-868A-56
7	237	13.0	309	1	US-08-518-878B-51
8	237	13.0	309	2	US-08-807-861A-51
9	237	13.0	309	3	US-08-470-868A-51
10	237	13.0	309	3	US-09-210-681-51
11	237	13.0	309	3	US-08-946-719A-51
12	236	12.9	320	2	US-08-933-750C-12
13	236	12.9	320	4	US-09-234-613-12
14	233.5	12.8	469	4	US-09-188-930-39
15	221.5	12.1	308	2	US-08-937-466-2
16	221.5	12.1	308	2	US-09-172-528-2
17	221.5	12.1	308	3	US-09-318-199-2
18	221.5	12.1	308	4	US-09-503-579-2
19	219.5	12.0	432	2	US-08-937-466-4
20	219.5	12.0	432	2	US-09-172-528-4
21	219.5	12.0	432	3	US-09-318-199-4
22	219.5	12.0	432	3	US-09-503-579-4
23	200	10.9	303	1	US-08-518-878B-37
24	200	10.9	303	1	US-08-294-522B-36
25	200	10.9	303	2	US-08-807-861A-37
26	200	10.9	303	2	US-08-470-868A-37
27	200	10.9	303	3	US-09-210-681-37

28 200 10.9 303 3 US-08-946-719A-37
29 198.5 10.9 306 5 PCT-US94-0979-1
30 197.5 10.8 312 4 US-09-142-565-2
31 194 10.6 307 2 US-08-807-861A-56
32 194 10.6 307 2 US-09-210-681-56
33 194 10.6 307 3 US-08-946-719A-56
34 192.5 10.5 311 2 US-08-775-009-32
35 185.5 10.1 298 3 US-08-961-871-10
36 181 9.9 256 2 US-08-937-466-6
37 181 9.9 256 2 US-09-172-528-6
38 181 9.9 256 3 US-09-318-199-6
39 181 9.9 256 4 US-09-503-579-6
40 177 9.7 311 2 US-08-775-009-33
41 125 6.8 312 4 US-09-188-930-142
42 90.5 5.0 3815 4 US-09-428-517-3
43 87 4.8 998 2 US-08-449-645A-20
44 87 4.8 998 2 US-08-702-367A-20
45 87 4.8 998 5 PCT-US95-04681-20

ALIGNMENTS

RESULT 1
US-08-933-750C-19
: Sequence 19, Application US/08933750C
: Patent No. 5932442
: GENERAL INFORMATION:
: APPLICANT: Lal, Preeti
: APPLICANT: Hillman, Jennifer L.
: APPLICANT: Bandman, Olga
: APPLICANT: Shah, Purvi
: APPLICANT: Au-Young, Janice
: APPLICANT: Yue, Henry
: APPLICANT: Guegler, Karl J.
: APPLICANT: Corley, Neil C.
: TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
: NUMBER OF SEQUENCES: 98
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/933 750C
: FILING DATE: September 23, 1997
: CLASSIFICATION: 536
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0356 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-855-0555
: TELEFAX: 415-845-4166
: TELEX:
: INFORMATION FOR SEQ ID NO: 19:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 351 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: SYNCOAT01

CLONE: 724157
US-08-933-750C-19

Query Match 100.0%; Score 1828; DB 2; Length 351;
Best Local Similarity 100.0%; Pred. No. 1,1e-179;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MADOPAGISPLQOMVAVSGTGAIVTSLFMTPLDVYKVRLOSQSPMASLSPSSRLMSLS 60
DB 1 MADOPAGISPLQOMVAVSGTGAIVTSLFMTPLDVYKVRLOSQSPMASLSPSSRLMSLS 60
QY 61 YTKMKCLLYCNGVLEPLYLCPNGARCATWFOPTFTGTMDAFVKIVRHEGRTTMSGLP 120
DB 61 YTKMKCLLYCNGVLEPLYLCPNGARCATWFOPTFTGTMDAFVKIVRHEGRTTMSGLP 120
QY 121 ATLVTVPATATFTAYDOLKAFICGRALTSPLYAPMVAGALARIGTVTVISPLELMTK 180
DB 121 ATLVTVPATATFTAYDOLKAFICGRALTSPLYAPMVAGALARIGTVTVISPLELMTK 180
QY 181 LOAHSVARELGACVRYTAAOGGWSLMLGWPALRDVPFSAIYWFNYELVKSMLNGLR 240
DB 181 LOAHSVARELGACVRYTAAOGGWSLMLGWPALRDVPFSAIYWFNYELVKSMLNGLR 240
QY 241 PKDQTSVGMSEFVAGISGTVAAVLTLPEDVYKTOBOVALGAEAVRVNPLHVDSTWLLLR 300
DB 241 PKDQTSVGMSEFVAGISGTVAAVLTLPEDVYKTOBOVALGAEAVRVNPLHVDSTWLLLR 300
QY 301 RIRASGTRKGLFAGFLPRITKAAPSCAIIISTYERKGSFFQRLNDRLLGG 351
DB 301 RIRASGTRKGLFAGFLPRITKAAPSCAIIISTYERKGSFFQRLNDRLLGG 351

RESULT 2

US-09-234-613-19
Sequence 19, Application US/09234613

Patent No. 6132973

GENERAL INFORMATION:

APPLICANT: Lal, Preeti

APPLICANT: Hillman, Jennifer L.

APPLICANT: Bandman, Olga

APPLICANT: Shah, Purvi

APPLICANT: Au-Yang, Janice

APPLICANT: Yue, Henry

APPLICANT: Guebler, Karl J.

APPLICANT: Corley, Neil C.

TITLE OF INVENTION: HUMAN REGULATORY MOLECULES

NUMBER OF SEQUENCES: 98

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/234,613

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/933,750

FILING DATE: September 23, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PR-0356 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 351 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: SYNCOAT01

CLONE: 724157
US-09-234-613-19

Query Match 100.0%; Score 1828; DB 4; Length 351;
Best Local Similarity 100.0%; Pred. No. 1,1e-179;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MADOPAGISPLQOMVAVSGTGAIVTSLFMTPLDVYKVRLOSQSPMASLSPSSRLMSLS 60
DB 1 MADOPAGISPLQOMVAVSGTGAIVTSLFMTPLDVYKVRLOSQSPMASLSPSSRLMSLS 60
QY 61 YTKMKCLLYCNGVLEPLYLCPNGARCATWFOPTFTGTMDAFVKIVRHEGRTTMSGLP 120
DB 61 YTKMKCLLYCNGVLEPLYLCPNGARCATWFOPTFTGTMDAFVKIVRHEGRTTMSGLP 120
QY 121 ATLVTVPATATFTAYDOLKAFICGRALTSPLYAPMVAGALARIGTVTVISPLELMTK 180
DB 121 ATLVTVPATATFTAYDOLKAFICGRALTSPLYAPMVAGALARIGTVTVISPLELMTK 180
QY 181 LOAHSVARELGACVRYTAAOGGWSLMLGWPALRDVPFSAIYWFNYELVKSMLNGLR 240
DB 181 LOAHSVARELGACVRYTAAOGGWSLMLGWPALRDVPFSAIYWFNYELVKSMLNGLR 240
QY 241 PKDQTSVGMSEFVAGISGTVAAVLTLPEDVYKTOBOVALGAEAVRVNPLHVDSTWLLLR 300
DB 241 PKDQTSVGMSEFVAGISGTVAAVLTLPEDVYKTOBOVALGAEAVRVNPLHVDSTWLLLR 300
QY 301 RIRASGTRKGLFAGFLPRITKAAPSCAIIISTYERKGSFFQRLNDRLLGG 351
DB 301 RIRASGTRKGLFAGFLPRITKAAPSCAIIISTYERKGSFFQRLNDRLLGG 351

RESULT 3

US-09-068-140A-15

Sequence 15, Application US/09068140A

Patent No. 6281409

GENERAL INFORMATION:

APPLICANT: Mary Rose Woodhead, Mark Andrew Taylor

APPLICANT: and Rex Michael Brennan

TITLE OF INVENTION: Blackcurrant Promoters and Genes

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: SmithKline Beecham Corporation

STREET: 709 Swedeland Road

CITY: King of Prussia

STATE: PA

COUNTRY: USA

ZIP: 19406-0939

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/068,140A

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/EP96/04807

FILING DATE: No. 6281409ember 4, 1996

ATTORNEY/AGENT INFORMATION:

NAME: Dinner, Data L.

REGISTRATION NUMBER: 33,680
REFERENCE/DOCKET NUMBER: C70237
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5017
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 328 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHEetical: YES
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Rides nigrum
STRAIN: Ben Alder
US-09-068-140A-15

Query Match 17.7%; Score 324; DB 4; Length 328;
Best Local Similarity 27.0%; Pred. No. 3.5e-25;
Matches 92; Conservative 65; Mismatches 128; Indels 56; Gaps 8;

OY 8 GISPLQWVAVSGTGAIVTSLFPTPLDVVKVRLQSORPSMASLMSRLMSLYTKMKCL 67
DB 33 GLQFQFMIAAGSIAGSIEHMAHYVDLTKTRIQ----- 65
OY 68 LYCNGLVEPLYLCPNGARCATWFODPTREFTGMDAFKIVRHEGRTLMGSLPATLVMTV 127
DB 66 -----GSCS-----AQSAGLRQALGSLKKEGPAGLYRGIGAMGLGAG 105
OY 128 PATATYFATYDQK-AFLCGRALTSDLAPMYAGALRLGTYTISPLELMTKQLQAHV 186
DB 106 PAHAAYFEVYECKETFESHGDPNSGAHA--VSGVFATVASDAVITPMDVVKQRLQSS 163
OY 187 SYRELGACVATVAAGGKRSMLGWPALRDVPFSALYWFYELVKSMNLGRPK--DQ 244
DB 164 PYKGVAVDVRRLVEEGIGAFYASRTTYVMNAPPTAVHFAFYETAKKGLLEVSEPTAND 223
OY 245 TSVGNSFVAGGISTVAAVLTLPFDVVKTORQVALGAEAVRNPLHYDSTWLLRLRIRA 304
DB 224 ENLVAHATAGAAAGALAAVTTPLDVKTKQLQ-CQVCGCDRFSSTIOD--VIGSIYK 279
OY 305 ESGTGLFAGFLPRTIKAPSCAISMISTYERKSFQRLNQ 345
DB 280 KNGYGLMRGWIPRLFHPAALICWSTYEASKTFEOKLINE 320

RESULT 4
US-09-068-140A-10
Sequence 10, Application US/09068140A
Patent No. 6281409

GENERAL INFORMATION:
APPLICANT: Mary Rose Woodhead, Mark Andrew Taylor
APPLICANT: and Rex Michael Brennan
TITLE OF INVENTION: Blackcurrant Promoters and Genes
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Smithline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/068.140A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP96/04807
FILING DATE: No. 6281409ember 4, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Dimer, Dara L.
REGISTRATION NUMBER: 33,680
REFERENCE/DOCKET NUMBER: C70237
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5017
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 289 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHEtical: YES
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Rides nigrum
STRAIN: Ben Alder
US-09-068-140A-10

Query Match 17.5%; Score 320; DB 4; Length 289;
Best Local Similarity 26.9%; Pred. No. 7.5e-25;
Matches 90; Conservative 64; Mismatches 124; Indels 56; Gaps 7;

OY 15 MVASGTGAIVTSLFPTPLDVVKVRLQSORPSMASLMSRLMSLYTKMKCLLYCNVL 74
DB 1 MIAGSIAGSIEHMAHYVDLTKTRIQ-----IGSCS----- 32
OY 75 EPLYLCPNGARCATWFODPTREFTGMDAFKIVRHEGRTLMGSLPATLVMTPATIYF 134
DB 33 -----AQSAGLRQALGSLKKEGPAGLYRGIGAMGLGAPRHAHYF 73
OY 135 TAYDQK-AFLCGRALTSDLAPMYAGALRLGTYTISPLELMTKQLQAHVSTRELGA 193
DB 74 SYEYCKETFESHGDPNSGAHA--VSGVFATVASDAVITPMDVVKQRLQSSPYKGVVD 131
OY 194 CVRTVAAGGKRSMLGWPALRDVPFSALYWFYELVKSMNLGRPK--DQTSVGSF 251
DB 132 CVRRLVEEGIGAFYASRTTYVMNAPPTAVHFAFYETAKKGLLEVSEPTANDLVHA 191
OY 252 VAGGISTVAAVLTLPFDVVKTORQVALGAEAVRNPLHYDSTWLLRLRAESGTGKL 311
DB 192 TAGAAGALAAVTTPLDVKTKQLQ-CQVCGCDRFSSTIOD--VIGSIYKKNYVGL 247
OY 312 FAGFLPRTIKAPSCAISMISTYERKSFQRLNQ 345
DB 248 MRGWIPRLFHPAALICWSTYEASKTFEOKLINE 281

RESULT 5
US-08-518-878B-56
Sequence 56, Application US/08518878B
Patent No. 5702802
GENERAL INFORMATION:
APPLICANT: Tartaglia, Louis A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TREATMENT OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York

COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/518,878B
FILING DATE: 23-AUG-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-036
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 56:
SEQUENCE CHARACTERISTICS:
LENGTH: 299 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
US-08-518-878B-56

Query Match 13.0%; Score 237; DB 1; Length 299;
Best Local Similarity 25.1%; Pred. No. 2.7e-16;
Matches 85; Conservative 43; Mismatches 132; Indels 78; Gaps 10;

QY 14 QWASGTGAVVTSLEMTPLDVVKVRLQSORPSMASSELMSSRLMSLYTKMKCLLYCNGV 73
DB 6 KRLGAGTACIADLTLPDFAKVRLOIGES----- 37
QY 74 LEPLVCPNGACATWFOPTPTFT-----GTMDAFKIVRHEGRTIMSGLPALIVMTV 127
DB 38 -----QGPVRAVSAOYRGVGTILMVTRGPRSLYNGLVAGLQROM 80
QY 128 PATAIYFAYDOLKAFLC-----GRAITSDLYAPMVAGALARIGTVVISPLEMRKLOA 183
DB 81 SFAVSRIGLIDYVKQFYTKGSEHASISGRLLAGSTTGALA-----VAVAOPTVYVVRQA 136
QY 184 QHVS-----YRELGACVRTAQAQGMRSIMLGWPTALRDVFPESALYWNVEYLVKSWLNG 238
DB 137 QARAGGRRYQSTVNAVKTIAREEGFRIGMKGTSNVANALVNCALVYDILDAL-- 194
QY 239 LRPKQOT-SVGNSTFAAGISGTVAAVLTLPDYVKTO-QVALGAMEAVRNPLHVDSTW 296
DB 195 LKANIMTDDLPCHTFSAGAGCTTVIASPVVDVVKTRYNSALGOYSSAG---HCALTM 250
QY 297 LLRRIRAESGTGKGFAGLPRIRIKAPSCAIMISTYE 334
DB 251 L-----QKSGPRAFYKGFMPFSLKLGSNWVMEVITYE 282

RESULT 6
US-08-470-868A-56
Sequence 56, Application US/08470868A
Patent No. 5861485
GENERAL INFORMATION:
APPLICANT: Tartaglia, Louis C.
TITLE OF INVENTION: Compositions and Methods for the
TREATMENT OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie and Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,868A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-0031-999
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864
TELEX: 66441 PENNIE
INFORMATION FOR SEQ ID NO: 56:
SEQUENCE CHARACTERISTICS:
LENGTH: 299 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
US-08-470-868A-56

Query Match 13.0%; Score 237; DB 2; Length 299;
Best Local Similarity 25.1%; Pred. No. 2.7e-16;
Matches 85; Conservative 43; Mismatches 132; Indels 78; Gaps 10;

QY 14 QWASGTGAVVTSLEMTPLDVVKVRLQSORPSMASSELMSSRLMSLYTKMKCLLYCNGV 73
DB 6 KRLGAGTACIADLTLPDFAKVRLOIGES----- 37
QY 74 LEPLVCPNGACATWFOPTPTFT-----GTMDAFKIVRHEGRTIMSGLPALIVMTV 127
DB 38 -----QGPVRAVSAOYRGVGTILMVTRGPRSLYNGLVAGLQROM 80
QY 128 PATAIYFAYDOLKAFLC-----GRAITSDLYAPMVAGALARIGTVVISPLEMRKLOA 183
DB 81 SFAVSRIGLIDYVKQFYTKGSEHASISGRLLAGSTTGALA-----VAVAOPTVYVVRQA 136
QY 184 QHVS-----YRELGACVRTAQAQGMRSIMLGWPTALRDVFPESALYWNVEYLVKSWLNG 238
DB 137 QARAGGRRYQSTVNAVKTIAREEGFRIGMKGTSNVANALVNCALVYDILDAL-- 194
QY 239 LRPKQOT-SVGNSTFAAGISGTVAAVLTLPDYVKTO-QVALGAMEAVRNPLHVDSTW 296
DB 195 LKANIMTDDLPCHTFSAGAGCTTVIASPVVDVVKTRYNSALGOYSSAG---HCALTM 250
QY 297 LLRRIRAESGTGKGFAGLPRIRIKAPSCAIMISTYE 334
DB 251 L-----QKSGPRAFYKGFMPFSLKLGSNWVMEVITYE 282

RESULT 7
US-08-518-878B-51
Sequence 51, Application US/08518878B
Patent No. 5702902
GENERAL INFORMATION:
APPLICANT: Tartaglia, Louis A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TREATMENT OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

RESULT 9
US-08-470-868A-51
Sequence 51, Application US/08470868A
Patent No. 5861485
GENERAL INFORMATION:
APPLICANT: Tartaglia, Louis C.
TITLE OF INVENTION: Compositions and Methods for the
TREATMENT OF OBESITY
TITLE OF INVENTION: Treatment of Body Weight Disorders, Including Obesity
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie and Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA

Page 6

MEDIMUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/210,661
 FILING DATE:
 CLASSIFICATION:
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: 08/807,861
 FILING DATE: 26-FEB-1997
 APPLICATION NUMBER: US 08/518,878
 FILING DATE: 23-AUG-1995
 APPLICATION NUMBER: US 08/470,866
 FILING DATE: 06-JUN-1995
 APPLICATION NUMBER: US 08/294,522
 FILING DATE: 23-AUG-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Coruzzi, Laura A.
 REGISTRATION NUMBER: 30,742
 REFERENCE/DOCKET NUMBER: 7853-066
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 869-9741/8864
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 51:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 309 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: unknown

```

Query Match      13.0%   Score 237; DB 3; Length 309;
Best Local Similarity 25.1%; Pred. No. 2.8e-16;
Matches    85; Conservative 43; Mismatches 132; Indels 78; Gaps 10.
QY 14 QMASTGAVVNSLEMTPLDVVKVFRLQSQRSMASELMPSSRLMSLSLYTKWCLLYCNGV 73
Db 16 KFLGAGAACIADLTLPFLDITAKVRLQIGGES----- 47
OY 74 LEPLYLCPNARCATWFODPTRPT-----GTMDAFVIYNHEGTRILMSGLPTLVNV 127
Db 48 -----QGVRATVSAYQRYGMGTILTWREGRSLTNGLVAGLOROM 90
OY 128 PATAIYTAVDOLKAFELC-----GRAFTSGLVPANAGALARIQTVYISPLEMRTKLQA 183
Db 91 SFASAIRIGLDYSKQFYTKESEHASIGSKLLAGSTTGALA----VAVAQPDVVKKVFEQA 146
OY 184 QHVS-----YRELACVRTAVAAGNRSIMLWGEPALRDVFFSKLYFNFIETVSMUNG 238
Db 147 QARAGGRRRGYSTVMNKVTIAREEGRGLMKTSPPVANAINATVNCALVELTYDDLKDAL-- 204
OY 239 LRPDQT-SYGMSPVAGIGSIGTAAVALTLTPEDVVKRO-QOVALLGAAMEAVRVNPFLHVDSIW 296
Db 205 LKANIMTDDEPCHTTSFAGFGCTTYIASVDVDVKTIRYNSKLGYOSSAG-----HCALTM 260
OY 297 LLRIIRAESGTEKGLFAGFLPRIIKAAAPCAIMISTYE 334
Db 261 L-----QKESPRAFYKGFMPSFLRLGSNMVAVFEITYE 292

```

RESULT 11
US-08-946-719A-51
; Sequence 51, Application US/08946719A
; Patent No. 6121017
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
;

```

: ADDRESSEE: Pennie & Edmonds LLP
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 10036-2711
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/946,719A
: FILING DATE: 8-OCT-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/807,861
: FILING DATE: 26-FEB-1997
: APPLICATION NUMBER: US 08/518,878
: FILING DATE: 23-AUG-1995
: APPLICATION NUMBER: US 08/470,868
: FILING DATE: 06-JUN-1995
: APPLICATION NUMBER: US 08/294,522
: FILING DATE: 23-AUG-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Coruzzi, Laura A.
: REGISTRATION NUMBER: 30,742
: REFERENCE/DOCKET NUMBER: 7853-107
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 790-9090
: TELEFAX: (212) 869-9741/8864
: TELEX: 66141 PENNIE
: INFORMATION FOR SEQ ID NO: 51:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 309 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: unknown
: US-08-946-719A-51

```

```

Query Match 13.0%; Score 237; DB 3; Length 309;
Best Local Similarity 25.1%; Pred. No. 2.8e-16;
Matches 85; Conservative 43; Mismatches 132; Indels 78; Gaps 10;

OY 14 OMVAGSTGAVVNTSLFMTPLDVVKVRLQSORPMSASELMPSSRLMSLSTKWKCLLYKGV 73
DB 16 KFLGAGTACINDLITFPLDTAKVKRLQIOGES----- 47
OY 74 LEPLVLCPRGACATWFOPTREF-----GTMDAFVKIVRHEGRTILMSGLPATLVMTV 127
DB 48 -----GQPVRAIVSAQYRGVNGTILFVFRPREGRSLYNGVLQGRM 90
OY 128 PATATATTAAYDOLKAFIC-----GRALTSPLAPMAGALARIGTYVISPLELMTKQIA 183
DB 91 SPASVRIGLYDVKFYTKGSEHASIGSRLLAGSTTGALA---VAVAPPTVVVVRQDA 146
OY 184 OHVS-----YRELGACVTRAVAGGWSRLMGPTALRDVPEFSALVWNTYELVSWLNG 238
DB 147 QARAGGRRYQSTVNNAKYTIAREEGFRGLMKSTSPRVARNALVNCALVYTLIDAL-- 204
OY 239 LRPKDOT-SVGSFVAGGISGTVAAVLTLPDFVKTQ--QVALGAMEAVRNPLHVDSTW 296
DB 205 LKANLMTDOLPCHFTSAFGAGCTTYIASPVVDVVKTRVWNSALQGVSSAG---HCALTM 260
OY 297 LLRLRRASGSKGLFAGLPLRIITKAAPSCAIMISTYE 334
DB 261 L-----QKEGPRAFYKGPSPFLRLGSLWNVVMEVYTE 292

```

RESULT 12
US-08-933-750C-12
Sequence 12, Application US/08933750C

```

: Patent No. 5932442
: GENERAL INFORMATION:
: APPLICANT: Lal, Preeti
: APPLICANT: Hillman, Jennifer L.
: APPLICANT: Bandman, Olga
: APPLICANT: Shah, Puri
: APPLICANT: Au-Young, Janice
: APPLICANT: Yue, Henry
: APPLICANT: Guegler, Karl J.
: APPLICANT: Corley, Neil C.
: TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
: NUMBER OF SEQUENCES: 98
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/933,750C
: FILING DATE: September 23, 1997
: CLASSIFICATION: 536
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0356 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-855-0555
: TELEFAX: 415-845-4166
: TELEX:
: INFORMATION FOR SEQ ID NO: 12:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 320 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: SPINNOT02
: CLONE: 207452
: US-08-933-750C-12

```

```

Query Match 12.9%; Score 236; DB 2; Length 320;
Best Local Similarity 22.0%; Pred. No. 3.7e-16;
Matches 83; Conservative 58; Mismatches 136; Indels 100; Gaps 12;

OY 3 DQDPAG--ISPLQOMVAGSTGAVVNTSLFMTPLDVVKVRLQSORPMSASELMPSSRLMSLS 60
DB 5 DEKPGRRNNTKQVAVASVSGIVTRALISPDFVYKIRFOLEHRLSR----- 52
OY 61 YTKMKCLLYKCNVLEPLVLCPRGACATWFOPT--RFTGMDAFVKIVRHEGRTILMSG 119
DB 53 -----SDPSAKYIGILQASQILOEBEPTLFMKGNH 82
OY 120 PATLVMTPATATTAAYDOLKA-----FLGRALTSPLAPMAGALA 163
DB 83 VPAQILSTIGYAVQPLSTLEMLTELVRGSSVYDAREFSVHFVCG-----GLAA 129
OY 164 RLGVTVYISPLELMTKQIAOHVS--YRELGACVTRAVAGGWSRLMGPTALRDVPEF 221
DB 130 CWATLTV--HPVDVLTTRPAAGSEPKVYNTLRLRAVGTWVRSEGPQVFGKGLAPTLAIPY 188
OY 222 SALVWFNVELVK---SWLNGLRPKDOTSVGMSFVAGGISGTVAAVLTLPDFDVVKTOROVA 278

```

Db 189 AGIOFSCYSLSKHLKYMALPAEGKKNENL-ONLLCGSGAGVISKTLTYPLDLFKRLQY- 246
 QY 279 LGAMEAVRNPLHVDSTWLLLRIRA-----ESGTGILFAGFLPRITKAAPSCAI 328
 Db 247 -GGFE-----HARAFQGVRRYKGLMDCAKOVLOKEGALGFPGKLSLSLKAALSTGE 298
 QY 329 MISTYEFKSFQRLNQ 345
 Db 299 MEFSTYEFQCNVPHCMNR 315

RESULT 13

US-09-234-613-12
 ; Sequence 12, Application US/09234613
 ; Patent No. 6132973
 ; GENERAL INFORMATION:
 ; APPLICANT: Lal, Preeti
 ; APPLICANT: Hillman, Jennifer L.
 ; APPLICANT: Bandman, Olga
 ; APPLICANT: Shah, Purvi
 ; APPLICANT: Au-Young, Janice
 ; APPLICANT: Yue, Henry
 ; APPLICANT: Guegler, Karl J.
 ; APPLICANT: Corley, Neil C.
 ; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
 ; NUMBER OF SEQUENCES: 98
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FASTSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/234,613
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/933,750
 ; FILING DATE: September 23, 1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Billings, Lucy J.
 ; REGISTRATION NUMBER: 36,749
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-855-0555
 ; TELEFAX: 415-845-4166
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 12:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 320 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: SPINNOT02
 ; CLONE: 207452
 ; US-09-234-613-12

Query Match 12.9%; Score 236; DB 4; Length 320;
 Best Local Similarity 22.0%; Pred. No. 3.7e-16;

Matches 83; Conservative 58; Mismatches 136; Indels 100; Gaps 12;

QY 3 DDPAG--ISPLQOMVAGSGAVVTSLEMTPLDVVKVRLQSGRPSMASELMPSSRLMSLS 60
 Db 5 DPKPGRRNNTKFOYAVVAGSVSGLVTRALISPPDVIKIRFOQHEHLSR----- 52

QY 61 YKMKCLLYCNGVLEPLYLCPNGARCATWFOPT-RTGTMDAFVKIVNHEGRTLSGL 119
 Db 53 -----SPPSAKHGIIQASRQILQEGPTAWKGH 82
 QY 120 PATLVNTVPATAIYPTAYDQKA-----FLGRLTSLYAPMAGALA 163
 Db 83 VPAQIISIGYGAOVFLSFMELTELVRHGSVYDAREFSVHFVCG-----GLAA 129
 QY 164 RLGTVTISPLELMTKRLQAOHVS--YRELGACVPTAANOOGMRSLMLGMPALRDVPE 221
 Db 130 CMATLTV-HPVDLRTKRFPAAGPEPKYNTLKNHAVGIMYSEGOVYFKGLAFTLLAIPFY 188
 QY 222 SALWPNVELVK---SWLNGLRPKDQTSVGSMSFVAGISGTYAAVLTLPFDVYKTRQROVA 278
 Db 189 AGIOFSCYSLSKHLKYMALPAEGKKNENL-ONLLCGSGAGVISKTLTYPLDLFKRLQY- 246
 QY 279 LGAMEAVRNPLHVDSTWLLLRIRA-----ESGTGILFAGFLPRITKAAPSCAI 328
 Db 247 -GGFE-----HARAFQGVRRYKGLMDCAKOVLOKEGALGFPGKLSLSLKAALSTGE 298
 QY 329 MISTYEFKSFQRLNQ 345
 Db 299 MEFSTYEFQCNVPHCMNR 315

RESULT 14

US-09-188-930-339
 ; Sequence 339, Application US/09188930A
 ; Patent No. 6150502
 ; GENERAL INFORMATION:
 ; APPLICANT: Watson, James D.
 ; APPLICANT: Strachan, Lorna
 ; APPLICANT: Sleeman, Matthew
 ; APPLICANT: Onrust, Rene
 ; APPLICANT: Murison, James Greg
 ; TITLE OF INVENTION: Compositions Isolated from Skin Cells
 ; FILE REFERENCE: 11000.1011c1
 ; CURRENT APPLICATION NUMBER: US/09/188,930A
 ; CURRENT FILING DATE: 1998-11-09
 ; NUMBER OF SEQ ID NOS: 348
 ; SOFTWARE: FASTSEQ for Windows Version 3.0
 ; SEQ ID NO 339
 ; LENGTH: 469
 ; TYPE: PRT
 ; ORGANISM: Mouse
 ; US-09-188-930-339

Query Match 12.8%; Score 233.5; DB 4; Length 469;
 Best Local Similarity 23.2%; Pred. No. 1.2e-15;
 Matches 77; Conservative 59; Mismatches 125; Indels 71; Gaps 9;

QY 13 QOMVAGSGAVVTSLEMTPLDVVKVRLQSGRPSMASELMPSSRLMSLSYTKKCLLYCNG 72
 Db 188 RHLVAGGAGAVSRTCTAPLRLKVLMO-----VHASSRNMM-----CI----- 226
 QY 73 VLEPLYLCPNGARCATWFOPTRTGTMDAFVKIVRHEGTRTMSGLPATLVNTPATAI 132
 Db 227 -----VGGFTQMRBEGAKSLMRGKNINVLKAPESAI 259
 QY 133 YFTAYDQKAEFLCGRALTSDLYAPMVAGALRLGTVTVISPLELMTKRLQAOHV-SYREL 191
 Db 260 KEMAVEQMKRLVGSQDETLLRIHERLVAAGSLAQAOSISYPMEEVLKTRMALRKTRGQSGM 319
 QY 192 GACVPTAANOOGMSLWGMGPTALRDVPEFSAIWFNVELVK-SWL-----NGLRPKDQT 245
 Db 320 LDCARRIILAKGVAAFYKGYIPNMGIIPYAGIDLAVETELKNTWLQRYAVNASDP---- 375
 QY 246 SVGMSFVAGISGTYAAVLTLPFDVYKTRQROVAL-----GAMDAVVRNPLHVDSTWLLLR 302
 Db 376 GVFTVILACGTISSTCGOLASTPLALVTRMQAQSIBGAPVYMS-----LKKQI 426

QY 303 RAESGTGGLFAGFLPRITKAPSCAIMISTYE 334
DB 427 LRTGAGGLYRGLAPNFKVTPAVSISYVYE 458

RESULT 15

US-08-937-466-2
Sequence 2, Application US/08937466
Patent No. 5846779
GENERAL INFORMATION:
APPLICANT: Zhang, Ning
APPLICANT: Amiral, M. Catherine
TITLE OF INVENTION: UCP3 Genes
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 75 DENISE DRIVE
CITY: HILLSBOROUGH
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/937,466
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: T97-009
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 308 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-937-466-2

Query Match 12.1%; Score 221.5; DB 2; Length 308;
Best Local Similarity 24.0%; Pred. No. 1.1e-14;
Matches 83; Conservative 49; Mismatches 147; Indels 67; Gaps 9;

QY 1 MADDOPAGISP--LOQMYASGTGAVVTSLEMTPLDVVKVRLQSQSPNASSELMPSRLMS 58
DB 1 MYGLQPSSEVPPTVYKFLGAGTACFADLTFPLDTAKVRLQIQENPQAQ----- 51
QY 59 LSYTRKMKCLLYCNGVLEPLVLCPNQARCATWFOPTFRGTMDAFVKIYRHGSTRITMSG 118
DB 52 -----SVQTRGVLTGTLTMTWRTGPRSPYSG 77
QY 119 LPATLVMTVPATAYFAVDOLKAFLCGR-ALTSDLXAPNVAGALARLGTVTISPLELM 177
DB 78 LVAGLHRQMSFASIRIGLYDSVKQFYTPKGAHSSVAIRILAGCTTGAMAVTCAQPTDV 137
QY 178 RTKLQAO-----HVSRELGACVFTAVVAVGGRSLMLGNGPTALRDVPFSALYWFNYE 230
DB 138 KVRQOAMIRLCTGGERKTRGTMDAYRTIAREGVRLMKGTWPNITRNAIVNCAEWYTD 197
QY 231 LVKSWLNGLRPKDOT-SVGNSEVAGGSGTVAVALTLPFDVYKQROVA-LGAEAVRVN 288
DB 198 IIKKEL--LBSHLEFTDNFPCHFVSAFGACATVVASPDVVVKTRYMNAPLGRYR-----S 251
QY 289 PLHVDSTWLLRLRRAESGTGKLEAGFLPRITKAPSCAIMISTYE 334

DB 252 PLH-----CMTKMAAQEGPTAFYKGFVPSFLRLGAMVNMMEVYIE 291

Search completed: April 19, 2002, 22:41:59
Job time: 3357 sec

Sat Apr 20 10:03:13 2002

us-09-840-787-19.ra1

Page 10

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2002, 20:34:46 ; Search time 172.76 Seconds
(without alignments)
8153.423 Million cell updates/sec

Title: US-09-840-787-68
Perfect score: 1643
Sequence: 1 GGGAGCGCTGACGACGACGAA.....GATGACTTCACAAAAA 1643

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: N_Geneseq_1101: *
2: /SIDS2/gcgdata/geneseq/geneseqn/NA1980.DAT: *
3: /SIDS2/gcgdata/geneseq/geneseqn/NA1981.DAT: *
4: /SIDS2/gcgdata/geneseq/geneseqn/NA1982.DAT: *
5: /SIDS2/gcgdata/geneseq/geneseqn/NA1983.DAT: *
6: /SIDS2/gcgdata/geneseq/geneseqn/NA1984.DAT: *
7: /SIDS2/gcgdata/geneseq/geneseqn/NA1985.DAT: *
8: /SIDS2/gcgdata/geneseq/geneseqn/NA1986.DAT: *
9: /SIDS2/gcgdata/geneseq/geneseqn/NA1987.DAT: *
10: /SIDS2/gcgdata/geneseq/geneseqn/NA1988.DAT: *
11: /SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT: *
12: /SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT: *
13: /SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT: *
14: /SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT: *
15: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT: *
16: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT: *
17: /SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT: *
18: /SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT: *
19: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT: *
20: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT: *
21: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT: *
22: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT: *
23: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1449	88.2	1560	21	AAC90452 Human uncoupling p
2	1405.4	85.5	1662	21	AAC76992 Human ORFX ORF2547
3	532	32.4	537	21	AA066605 Human immunogenic
4	532	32.4	537	22	AAH93722 Human prostate-spe
5	532	32.4	537	22	AAH85036 Human prostate-spe
6	532	32.4	537	22	AAH02787 Prostate tumour an
c	7	243.8	14.8	7112	21 AAC76691 Human ORFX ORF2246
8	168.4	10.2	174	16	AAT26018 Human gene signatu
9	144.2	8.8	397	21	AAH30252 Human colon cancer
c	10	74.8	4.6	452	22 AA130053 Probe #7739 used t
11	74.8	4.6	496	22	AA139123 Probe #7809 used t

c	12	65.4	4.0	72604	20	AA210752	Genomic sequence o
c	13	64.2	3.9	4231	21	AAC77105	Human ORFX ORF2660
c	14	63.6	3.9	444	22	AAE65290	Novel human polyu
c	15	63.4	3.9	546	22	AAH09622	Human cDNA clone (
c	16	63.4	3.9	1517	22	AAH17208	Human cDNA sequenc
c	17	63	3.8	52216	22	AAH28355	Nucleotide sequenc
c	18	62.8	3.8	4089	20	AAZ08807	Human cerebral pro
c	19	62.8	3.8	87350	18	AAH83003	Human WRN genomic
c	20	62	3.8	5009	19	AAV65634	First EcorI nucleoe
c	21	62	3.8	26928	20	AAZ32184	Human prothrombin
c	22	62	3.8	48000	22	AAE27996	Human calcium sens
c	23	61.8	3.8	2492	21	AAC78199	Human cancer assoc
c	24	61.6	3.7	700	22	AAH92306	Human inflammatory
c	25	61.6	3.7	1013	21	AAC68083	Human secreted pro
c	26	61.6	3.7	7680	21	AAH89439	Human kidney relat
c	27	61.6	3.7	9439	22	AAE33723	Human kidney relat
c	28	61.6	3.7	50000	21	AAH6363	Polymorphic repeat
c	29	61.6	3.7	97662	22	AAE83908	Genomic sequence o
c	30	61.6	3.7	106746	21	AAH10225	Human PCNA-1 genom
c	31	61.4	3.7	215	21	AAC29167	Human secreted pro
c	32	61.4	3.7	1545	21	AAC59037	Human secreted pro
c	33	61.4	3.7	7680	21	AAH89439	Human secreted pro
c	34	61.2	3.7	2202	22	AAH17401	Human cDNA sequenc
c	35	61.2	3.7	17858	22	AAE4114	Human bladder rela
c	36	61.2	3.7	160552	22	AAE02697	Human glycosyl sul
c	37	61	3.7	323	21	AAC07236	Human secreted pro
c	38	61	3.7	997	21	AAZ61628	CDNA encoding murI
c	39	61	3.7	997	22	AAC99561	SKIN cell CDNA, SE
c	40	61	3.7	1808	22	AAC99806	SKIN cell CDNA, SE
c	41	61	3.7	1816	21	AAZ61789	CDNA encoding murI
c	42	61	3.7	1816	22	AAC99722	SKIN cell CDNA, SE
c	43	61	3.7	2270	22	AAH16683	Human cDNA sequenc
c	44	61	3.7	110000	22	AAE84800	Nucleotide sequenc
c	45	61	3.7	161425	22	AAH02340	Human ARAPI0 gene

ALIGNMENTS

RESULT 1	AAC90452 standard; cDNA: 1560 BP.
ID	AAC90452
AC	AAC90452;
XX	
DT	12-MAR-2001 (first entry)
XX	
DE	Human uncoupling protein cDNA #1.
XX	
KW	Human; uncoupling protein; immunosuppressive; antiarthritic;
KW	antirheumatic; antiproliferative; cardiant; vasotropic;
KW	cerebroprotective; neuroprotective, antibacterial; ophthalmological;
KW	gastrointestinal; nephrotropic; gynaecological; vulvneary; thrombotic;
KW	gene therapy; cancer; wound; infectious disease; thrombosis; arthritis;
KW	Intertility; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200061614-A2.
XX	
PD	19-OCT-2000.
XX	
PF	06-APR-2000; 2000WO-US09534.
XX	
PR	09-APR-1999; 99US-0128701.
PR	08-JUL-1999; 99US-0142821.
PR	18-AUG-1999; 99US-0149448.
XX	12-NOV-1999; 99US-0164751.
XX	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Ruben SM, Nt J, Komatsoulis G, Rosen CA, Soppet DR;

DR WPI: 2000-656322/63.
DR P-PSDB: AAB50378.
XX Uncoupling proteins and nucleic acid sequences encoding them, useful
PT for detecting, preventing and treating proliferative, neurological,
PT immune system, cardiovascular and gastrointestinal disorders -
XX
XX Claim 1; Page 303; 343pp; English.
XX
CC The present sequence is one of eighteen isolated nucleotide sequences
CC encoding uncoupling proteins. The nucleotide sequences may be used for
CC the detection of various disorders such as cancer, for chromosome
CC identification, as chromosome markers and for numerous other diagnostic
CC or research purposes. The uncoupling protein encoded by the nucleotide
CC sequences may be used to treat disorders such as neural, immune,
CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
CC renal and proliferative disorders, wounds, infectious diseases,
CC thrombosis, arthritis, and infertility.
CC
SO Sequence 1560 BP; 286 A; 492 C; 459 G; 323 T; 0 other:

Query Match 88.2%; Score 1449; DB 21; Length 1560;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1452; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 187 CAGACGTAAAGCTTCAAGATGGCTGACGACGACCTGGGGCATGACCCCTCCAGCAA 246
DB 102 cagagctgaagcttcaagatggtgacagagcctgaggatagagccctccagcaa 161
QY 247 ATGGTGGCTTCAAGACCGGGGCTGTGTACTCTCTCTTATACACACCTTGACGTG 306
DB 162 atggtggcttcaagacccgggtgtgtactctcttcaatgaccccttgagctg 221
QY 307 GTGAGGTTCGGCTCAGCTTCAGCGGCCCTTCATGGCGAGAGCTGATGCTTCCTCC 366
DB 222 gtgaaagtgtgcgtgcagatccagcgcctccatgagcagagctgtagcttccccc 281
QY 367 AGACTGTGAGAGCCTCTCTTATACCAATGGAAGTGCCTCTGTATGCAATGGTCTCTG 426
DB 282 agactgtgagagcctctctctatacaaatggaagtgcctctgtatctgcaatgtgtctg 341
QY 427 GAGCTCTGTACTCTGTGCCAAATGATGTCGCTGTGCACCTGGTTTCAGACCTTAC 486
DB 342 gagctctgtactctgtgccaaatgattgcgcgtgtgcacactggtttcaagaccctac 401
QY 487 CGCTCAGTGGACCATGATGCTTCGTAAGATCGTAGAGCAGAGGACACGAGAC 546
DB 402 cgcttcactgacacatgtagtgccttctgtagatcgtgagcagagagggccagagacc 461
QY 547 CTCTGGAGCGGCTCCCGGACCCCTGTGTGATGACTGTGACGACTACCGCATTTACTTC 606
DB 462 ctctggagcggcctcccgacacctgtgtgactgtgacgactacacgcaatctacttc 521
QY 607 ACTGCTTATGACCAACTGAAGCCTTCTGTGTGTGTCAGGCCCTTACTCTGACTTAC 666
DB 522 actgcttattgacaaactgaagccttctgtgtgtgtcagaccctgaacctgaacctac 581
QY 667 GCACCATGATGGCTGGGCGCTGGCCGCTGGACCGAGATGATGATGATGATGATGATG 726
DB 582 gcaaccatgattggctgggctggccgctggacccgagatgattgattgattgattgatt 641
QY 727 GAGCTTATCGGACAAAGCTTCAGGCTCAGCATGTGTGTACCGGAGAGCTGGTCTGT 786
DB 642 gagcttattcggaacaaagcttcagagctcagcatgtgtacccggagctgtgtcctgt 701
QY 787 GTTGAATTCAGATGCTCAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 846
DB 702 gtctgaatttcagatgctcagagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 761
QY 847 GCCCTTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 906
DB 762 gcccttcagatgattgattgattgattgattgattgattgattgattgattgattgatt 821

QY 907 TGGCTCAATGGGCTCAGCGGAGGACGAGCTTGTGTGGCATGAGCTTTGTGGCTGT 966
DB 822 tggctcaatgggctcagcgagagacgagcttgtgtggcatgagctttgtggctgt 881
QY 967 GGCATCTCAGGAGCGGTGGCTGACGTGACTTACCTTTGAGTGTGTAAGACCCAA 1026
DB 882 ggcattctcagagcggtggctgacgtgacttaccctttgagtggttaagacccaa 941
QY 1027 CCGCAGTGTGCTGTGGAGCAGATGAGGCTGTGAGATGAACCCCTGATGTGACTTC 1086
DB 942 ccgcagtgctgtgagcgagatgagctgtgagatgagaaacccctgattgtgacttc 1001
QY 1087 ACCGTGCTGCTCTCGGAGAGATCCGGGCGAGTGGGCGACCAAGAGACTTTTGACAG 1146
DB 1002 accgtgctgctctcggagagatccgggcgagtgggaccaaagagacttttgacag 1061
QY 1147 TTCCTTCCCGGATCATCAAGGCTGCCCTTCCTGCTGCATGATCATGACGACATGAG 1206
DB 1062 ttcccttcccgatcatcaaggctgcccttctgtgcataatgacgacccatgag 1121
QY 1207 TTGGCAAAAGCTTCTTCAGAGGCTGAACAGACCGGCTTCTGGGCGCTGAAGGGG 1266
DB 1122 ttggcaaaagcttcttcagagctgaacacggcttctgggctgaaggagg 1181
QY 1267 CAGAGAGCGCAGACCCGCTCTCTCCACGATGGGAGAGGGCAAGAGAGACCCAGCC 1326
DB 1182 caagagagcgaccccgctctctccacgattgggagagggcagagagaccagcc 1241
QY 1327 AAGTGCCTTTTCTTCAGACATGAGGAGGGGCTGTTCCTTCCTCCCGGCGACGAC 1386
DB 1242 aagtgccttttcttcagacatgagggagggctgttccttccctcccgagcaag 1301
QY 1387 CTCACAGGAGGCGCTGCTCCTCTGGGGCGCCAGACACTTCTCAGACAACTTTCTCT 1446
DB 1302 ctcaagagagggctgctcctctggggcgccagacacttctcagacaaactttctct 1361
QY 1447 GCTGCTCAGTGTGGGATATCATCTTACCCACCCCCCAAGTTCAAGACCAATCTTCC 1506
DB 1362 gctgctcagtggtggatatacttaccacccccaagtccaagacaaactcttc 1421
QY 1507 AGCTGCCCCCTCGTGTTCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1566
DB 1422 agctgccccctcgtgttccctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1481
QY 1567 AAGCCTCAGGCTGTGTGATGCTCCCTGACCTTGTAACTTAACTTAAAGATGAT 1626
DB 1482 aagcctcagcctgtgtgattcctcctgaacctgttaattcctaagtaagatgat 1541
QY 1627 GAATTCAAAAAAA 1643
DB 1542 gaataaaaaa 1558

RESULT 2
AAC76992
ID AAC76992 standard; cDNA; 1662 BP.
XX
XX AAC76992;
AC
AC
DT 08-FEB-2001 (first entry)
XX
XX Human ORFX ORF2547 polynucleotide sequence SEQ ID NO:5093.
DE
DE Human; open reading frame; ORFX; detection; cytosolic; hepatotropic;
KW vulnerability; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritis; immunosuppressant; cardiast;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;

KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive; ss.
 XX Homo sapiens.
 OS
 PN MO200058473-A2.
 PD 05-OCT-2000.
 XX
 XX 31-MAR-2000; 2000WO-US08621.
 XX
 XX 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 XX Shmukets RA, Leach M;
 PI
 DR WPI: 2000-602362/57.
 DR P-PSDB: AAB42783.
 XX
 PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 XX Claim 5: Page 4281-4282; 5507pp; English.
 XX
 XX AACT74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORF open reading frames 1 to 3161. The ORF
 CC sequences have activities such as: cytoskeletal; hepatotropic; vulnary;
 CC antiproliferic; antiparkinsonian; nootropic; immunosuppressive;
 CC osteoplastic; anticonvulsant; antiallergic; immunosuppressive;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antiinflammatory; antibacterial; antiviral; antifungal; antineumatic;
 CC antihydroid; and antianemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORF-associated disorder. The
 CC nucleic acids can be used to express ORF proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 XX
 XX Sequence 1662 BP; 336 A; 516 C; 468 G; 341 T; 1 other;
 SQ

Query Match 85.5%; Score 1405.4; DB 21; Length 1662;
 Best Local Similarity 97.2%; Pred. No. 0;
 Matches 168; Conservative 0; Mismatches 1; Indels 42; Gaps 2;

QY 175 AGCTTGGTTTCAGAGCTGAAGCTTCAAGATGGCTGACAGACCGCTGGGCAATCAC 234
 Db 106 aggttggcttccaggaagctcgaagatgctgacacgaagaccctgcgggcatcagc 165
 QY 235 CCCCTCACAAATGGTGCTCAGGACACGGGGCTGTGGTTACCTCTCTTCATGACA 294
 Db 166 cccctccagcaaatggtgcctcagggcaccggggtctgtgttaacctcttcatgaca 225
 QY 295 CCCCTGAGAGTGTGATGATGGCTGCTGAGTCTCAGGCGCCCTCCATGGCCAGAGCTG 354
 Db 226 cccctgagagctgtgaaagctcgcctgcagctcagcgccctccatgacgagcgtg 285

QY 355 ATGCTTCTCTCCAGACTGTGGAG-----CCTCTCTATAC 390
 Db 286 atgcttctctccagactgtgagcctctctctataacaaatgcctcctctccataacc 345
 QY 391 AAATGGAAGTGCCTCTCTATTGGCAATGTGTCTCGAGAGCTCTTACTGTGCCCAAT 450
 Db 346 aaatggaagtgcctctctctatgtcaatgtgtctcggagccctcttactgtgccaaat 405
 QY 451 GGTGCCGCTGTGGCACTGGTTTCAGACCCCTTACCGCTTCTACTGTGGACATGTAGTGC 510
 Db 406 ggtgccgctgtgccaactgtgttccaagacctaccgcttcaactgtgaacatgtatgcc 465
 QY 511 TTCTGTAAGATCGTAGGCACAGAGGACACAGACCCCTCTGGAGCGGCTCCCGCCAC 570
 Db 466 ttctgtaagatcgttaggcacagagggcaccagacctcttgagggctctcccgccacc 525
 QY 571 CTGTGATGACTGTGCCAGACTACCGCATCTTACTTCACTGCTTATGACCACTGAAGGC 630
 Db 526 ctgtgatagtactgtgccagctaccgcatctacttcaactgtcctatgaccaaactgaagcc 585
 QY 631 TTCTGTGTGTGTGAGACCCCTGACCTCTGACCTTACCGACCCATGTGTGCTGGCGCTG 690
 Db 586 ttctgtgtgtgtgagacccctgacctctgacctctacgacccaatgtgtgtgcgtgctg 645
 QY 691 GCCCGCTTGGGACCGCTGACTGTGATCAAGCCCTTGAGACTTATGCGCAAGAGCTGCAG 750
 Db 646 gcccgcttgggacccgctgactgtatcaagccccctggagcttatgcgcaaaagctgcag 705
 QY 751 GCTCAGCATGTGTCTGTACCGGAGAGCTGGGTGCTGTTCGAACATGCAATGCTCAGGT 810
 Db 706 gctcagcatgtgtctgtaccggagagctgtgtcgtgtgtcgaacgtcagtgctcaaggt 765
 QY 811 GCGTGGCGCTCACTGTGGCTGGGGCTGGGGCCCGCTTGCAGATGTGCCCTTCTTC- 869
 Db 766 gctgtggcgctcaactgt 825
 QY 870 -----ACCGTGTACTGTTCAACTATGACTGTGTGAAGAGCTGGCTC 912
 Db 826 gtcatcccccaccccaagccctgtactgtgttcaactatgagctgtgtgaagctgtgctc 885
 QY 913 AATGGCTTACGCGCCAGAGACAGCACTTCTGTGGCATGATGCTTGTGGCTGTGGCAATC 972
 Db 886 aatggcttaccgctcgaagagacacacttctgtggcatgactgtgtgtgtgtgtgtgtgt 945
 QY 973 TCAGGAGAGTGTGAGTGTGACTTACCTTGTGAGTGTAAAGCCCAAGCCGAC 1032
 Db 946 tcaaggagctgt 1005
 QY 1033 GTGCTGTGGAGCGATGAGAGCTGTGAGAGTGAACCCCTGCATGTGGACTCCACTTG 1092
 Db 1006 gtgcttgtggagcgatgtgaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1065
 QY 1093 CTGTGCTGCGAGGATCCGGGCGCGAGTGGGCGACCAAGGAGCTTTTGACGCTTCTT 1152
 Db 1066 ctgtgctgcggagatccgggcgagctgcggcgacaaaggagctcttgcagcttctt 1125
 QY 1153 CCTCGATCATTAAGGCTGCCCTCTCTGTGCAATCATATATAGACCACTATGATTTGCG 1212
 Db 1126 cctcgatcataaagctgtgccccctctgtgtcatcatagtatgaacctatgaagctcggc 1185
 QY 1213 AAAAGCTTCTTCAGAGCGTGAACAGGACCGGCTTGTGGGGGCTGAAGAGGCAAGGA 1272
 Db 1186 aaaagcttcttccaagaagctgaacccaagacggcttctgtggggtctgtaaaagggcaga 1245
 QY 1273 GGCAAGAGCCCGCTCTTCCACGAGATGGGAGAGGAGGAGAGACCAACCAAGTGC 1332
 Db 1246 ggcagaagacccgctcttccacagatgggagggcgagggagggagaccagcaagatgc 1305
 QY 1333 CTTTTCCTCAGACATGAGGAGAGGGGCTGTTCCTTCTCCCTCCGGCCACAAGCTCCAG 1392
 Db 1306 ctttctcctaagacatgagggaggggtgtgttcttccctccctccggcgagcaagctccag 1365
 QY 1393 GGCAGAGGCTGTCCCTGTGGGCGGCCAGCATCTTCTCAGACCAACTTCTCTGTGCT 1452

```
|||||
Db 1366 ggcagagcgtccctcctggtggcgccacgacctccctcagacaactcttcctcgtgct 1425
OY 1453 CCAGTCGTGGGATATCATCTACCCACCCCAAGTTCAACACCAATCTCCAGCTGC 1512
Db 1426 ccagtcgtggtgatatcatctaccaccccccaagttcaagaacaaatcttcagctgc 1485
OY 1513 CCCCTTCGTGTTCCCTGTTGCTGTAGTGGGCATGTCTCAGAGCAAGAACGCC 1572
Db 1486 cccctctgttcttccctgtgttctgtcagctgggcaatctccaggaaccaagaagccc 1545
OY 1573 TCAGCCGTGTGTAGTCTCCCTGACCCCTGTGTAATTCCTTAAGCTAAAGATGAACTT 1632
Db 1546 tcagcctgtgtgagctccctcagacctgttaattccttaagtaagatgatgaactt 1605
OY 1633 CAAAAAATAA 1643
Db 1606 caaaaaaaa 1616

RESULT 3
AAA06606
ID AAA06606 standard; cDNA: 537 BP.
XX
AC AAA06606;
XX
DT 13-JUN-2000 (first entry)
XX
DE Human immunogenic prostate tumour protein cDNA sequence SEQ ID NO:387.
XX
KW Human; prostate cancer; diagnosis; tumour; gene therapy; detection;
KW immunogenic; cytosolic; vaccine; ss.
XX
OS Homo sapiens.
XX
PN WO200004149-A2.
XX
PD 27-JAN-2000.
XX
PF 14-JUL-1999; 99WO-US15838.
XX
PR 14-JUL-1998; 98US-0115453.
PR 14-JUL-1998; 98US-0116134.
PR 23-SEP-1998; 98US-0159812.
PR 23-SEP-1998; 98US-0159822.
PR 15-JAN-1999; 99US-0232149.
PR 15-JAN-1999; 99US-0232880.
PR 09-APR-1999; 99US-0288946.
XX
FA (CORI-) CORIXA CORP.
XX
PI Dillon DC, Harlocker SL, Yugu J, Xu J, Mitcham JL;
XX
WP: 2000-171268/15.
XX
DR New polypeptide useful for treating and diagnosing prostate cancer
PT comprises an immunogenic portion of prostate tumor protein -
XX
PS Claim 50; Page 235-236; 263pp; English.
XX
XX The present invention describes isolated polypeptides, comprising an
CC immunogenic portion of a prostate tumour protein (PTP). The polypeptides
CC and polynucleotides encoding them have cytosolic activity and can be
CC used in vaccines and in gene therapy. The polypeptides and
CC polynucleotides encoding them, antigen presenting cells which express
CC the polypeptides, antibodies against the polypeptides and vaccines
CC comprising them can be used for inhibiting the development of prostate
CC cancer in a patient. The polypeptides can be used to generate antibodies
CC or anti-idiotypic antibodies for passive immuno therapy. A portion of
CC the polynucleotides encoding the polypeptides can be used as a probe or
CC to modulate the expression of the polypeptides. AAA06241 to AAA06691 and
CC AA82000 to AA82020 represent sequences used in the exemplification of
CC the present invention.
```

```
XX
SQ Sequence 537 BP; 116 A; 167 C; 138 G; 116 T; 0 other;
Query Match 32.4%; Score 532; DB 21; Length 537;
Best Local Similarity 100.0%; Pred. No. 6,7e-125;
Matches 532; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1112 GGGCGAGTCGGGACCAAGGAGACCTTTGACAGGCTTCTCTCGGATCAACAGGCTG 1171
Db 1 999ccgagtcgagacacaaaggactcttgcaggtcttcccttcgacatcaaaagctcg 60
OY 1172 CCCCTCTCTGTGCATCATGATCAGACACCTATGATTCGGCAAAAGCTTCTTCAGAGCG 1231
Db 61 cccctcctctgtgcatcatcatcagacactatgagttcggcaaaagctcttccagagcg 120
OY 1232 TGAACGAGACCGGCTTCTGGCGGCTGAAGAGGCGCAAGAGCAAGACCCGCTCTTC 1291
Db 121 tgaaccagaccggtctctcgtgaggaagggaaggaaggaaggaaggaaggaaggaagga 180
OY 1292 CCACGATGGGAGAGGAGGAGAGAGAGACCAAGCCCAAGTCCCTTTCACAGACCTGAGG 1351
Db 181 ccaaggatgaggagaggaaggaggaaggaaggaaggaaggaaggaaggaaggaaggaagga 240
OY 1352 GAGGGGGCTTGTTCCTTCCTCCCTCCGCGACAGCTCCAGGCGAGGGCTGCTCCCTGG 1411
Db 241 gagggggctgttcccttccctccctcccgagacagctccagggcagggctgtccctctcg 300
OY 1412 GCGGCCAGACACTTCTCTCAGACACAATTTCTTCTGCTGCTCAGTCTGGGGATCATCA 1471
Db 301 gggccagacacttctcagacacacttctcctgtcagtcgtygggagatcata 360
OY 1472 CTTACCCACCCCAAGTTCAAGACCAATCTTCAGCTGCCCTCGCTGTTCCCTGT 1531
Db 361 cttaaccacccccaagttcaagacaaactcttcaagctgtccctctgttccctcgt 420-
OY 1532 GTTTCCTGTAGCTGGGATGCTCTCCAGAACCAAGACCTTCAGCTGCTGTCTCC 1591
Db 421 gtttctgtagctgggacatgtctcaggaaccaagaagccctcagcgtgtgtagcttcc 480
OY 1592 CTGACCCCTGTGTAATTCCTTAAGTCTAAAGATGATGACTTCAAAAAA 1643
Db 481 ctgacctgttaattccttaagtaagatgacttaaaaaa 532

RESULT 4
AAH93722
ID AAH93722 standard; cDNA: 537 BP.
XX
AC AAH93722;
XX
DT 04-OCT-2001 (first entry)
XX
DE Human prostate-specific cDNA sequence CGI-69.
XX
KW Human; prostate cancer; prostate-specific; diagnosis; vaccine;
KW cytosolic; gene therapy; metastasis; ss.
XX
OS Homo sapiens.
XX
PN WO200151633-A2.
XX
PD 19-JUL-2001.
XX
PF 16-JAN-2001; 2001WO-US01574.
XX
PR 14-JAN-2000; 2000US-0483672.
XX
FA (CORI-) CORIXA CORP.
XX
PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
PI Kuo J, Kuo J, Fanger GR, Day CH, Retter MM, Stolk JA, Skelky YAM;
PI Wang A, Weagher MT;
```

XX WP1: 2001-425873/45.
 XX
 XX New polynucleotide encoding a prostate-specific protein, for
 PT diagnosing, monitoring and treating prostate cancer in a patient and
 PT for use in vaccines -
 XX
 XX Claim 1; Page 359-360; 543pp; English.
 XX
 CC The present invention describes polynucleotide sequences (I) which encode
 CC prostate-specific proteins (II). (I) and (II) have cytostatic activity,
 CC and can be used in vaccine production and gene therapy. (I), (II),
 CC antibodies to (II), fusion proteins comprising (II), and isolated
 CC T cells prepared using (I) or (II) are used to treat cancer in a patient.
 CC (I) and the antibodies are also used in the detection of cancer in a
 CC patient. The cancer that is diagnosed or treated is particularly
 CC prostate cancer. (I) and (II) can be used in vaccines. The antibodies or
 CC (I) can be used for monitoring the progression of cancer in a patient.
 CC (I) and (II) can also be used to improve diagnostic and therapeutic
 CC methods for prostate cancer. They can indicate the level of metastasis
 CC as well as the prostate volume. AAH9357 to AAH9394 and AAH0115 to
 CC AAH0138 represent polynucleotide and amino acid sequences used in the
 CC exemplification of the present invention.

XX Sequence 537 BP; 116 A; 167 C; 138 G; 116 T; 0 other;

Query Match 32.4%; Score 532; DB 22; Length 537;

Best Local Similarity 100.0%; Pred. No. 6.7e-125; Mismatches 0; Indels 0; Gaps 0;

XX 1112 GGGCCGATCGGGCCACCAAGGACTCTTGGAGGCTTCTCGATCATCAGGCTG 1171
 Db 1 99gcgagctcgggcacaaggagactcttgcaaggtcttcctcctcgtatcataagctg 60
 QY 1172 CCCCCCTCTGTCATCATGATCAGACCTATGATTTGGCAAAAGCTTCTTCAGAGC 1231
 Db 61 cccctctctgcatcatgatacagacattgattctgcaaaagctctctccagagc 120
 QY 1232 TGAACCAAGACCGGCTTCTGGGCGGCAAGGAGGAGGCAAGACCCGCTCTC 1291
 Db 121 tgaaccagagacggtctctggtcggtgaaaggagagagagaccccgctctc 180
 QY 1292 CCACGGATGGGAGAGGCGCAGAGAGACCCAGCCAGTGCCTTTCTTCAGACTGAG 1351
 Db 181 ccacggaatgggagagggcagagagagagccagccaagtgccttctcagactgag 240
 QY 1352 GAGGGGCTTGTTCCTTCCCTCCCTCCCGGACAGCTCCAGGGAGGCTGCTCTGG 1411
 Db 241 gaggggggtgttctccctccctcccgagacaagctccagggcagggctgtccctc 300
 QY 1412 GCGGCGCCAGCACTTCTCCAGACACAACTTCTCGTGTCTCAGTCTGGGATCATCA 1471
 Db 301 gcgggccagacactctcctcagacacaactctctcgtcccggtcggtggatcatca 360
 QY 1472 CTTACCAACCCCAAGATTCAAGACCAAAATCTTCCAGCTGCCCCCTTGTTCCTGT 1531
 Db 361 ctaccacaccccccagaagttcaagaacaaatctccagctgcccccttggttccctc 420
 QY 1532 GTTGTGCTAGCTGGGAGTGTCTCAGAGACCAAGAGCCCTTCAGCTCGGTGTAGTCC 1591
 Db 421 gttgtgctagctgggagtgctccaggaagaccagaagccctcagccgtggtgtctcc 480
 QY 1592 CTGACCCCTGTGTAATCTCTTAAGTCTAAAGATGATGAATTCAGAAATTAATAA 1643
 Db 481 ctgacccctgttaattccttaagctaaagatgagtgaactctcaaaaaa 532

RESULT 5
 AAH85036
 ID AAH85036 standard; cDNA; 537 BP.
 XX
 XX AAH85036;

XX 25-SEP-2001 (first entry)
 DT
 XX Human prostate-specific cDNA sequence CGI-69.
 DE
 XX Human prostate cancer; therapy; diagnosis; cat eye syndrome;
 KW chromosome 22q11.2; prostate-specific protein; chromosome 1;
 KW prostate specific antigen; PSA; ss.
 XX

XX Homo sapiens.

XX WO200134802-A2.

XX 17-MAY-2001.

XX 09-NOV-2000; 2000WO-US30904.

XX 12-NOV-1999; 99US-0439313.

XX 18-NOV-1999; 99US-0443686.

XX (CORI-) CORIYA CORP.

XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
 PI Kalos MD, Retter MW, Stolk JA, Day CH, Skelky YAW, Wang A;
 XX WP1: 2001-308785/32.

XX Isolated polypeptide comprising at least an immunogenic portion of a
 PT prostate-specific protein, useful in the diagnosis and therapy of a
 PT prostate cancer -
 XX

XX Claim 31; Page 259; 325pp; English.

XX The present invention describes an isolated polypeptide (P1) comprising
 CC at least an immunogenic portion of a prostate-specific protein, or its
 CC variant. Also described are polynucleotides (N1) encoding (P1). (P1) and
 CC (N1) have cytostatic activity and can be used in vaccine production.
 CC The polypeptides, nucleic acids and antibodies from the present
 CC invention are useful in the diagnosis and therapy of prostate cancer.
 CC Prostate specific genes P704P, P712P, P774P, P775P and B305D are located
 CC in a genomic region on chromosome 22q11.2 known as the Cat Eye Syndrome
 CC region. Prostate specific antigen (PSA) P501S was located on
 CC chromosome 1. AAH84671 to AAH85143 and AAG99000 to AAG99077 represent
 CC polynucleotide and polypeptide sequences used in the exemplification
 CC of the present invention.

XX Sequence 537 BP; 116 A; 167 C; 138 G; 116 T; 0 other;

Query Match 32.4%; Score 532; DB 22; Length 537;

Best Local Similarity 100.0%; Pred. No. 6.7e-125; Mismatches 0; Indels 0; Gaps 0;

XX 1112 GGGCCGATCGGGCCACCAAGGACTCTTGGAGGCTTCTCGATCATCAGGCTG 1171
 Db 1 99gcgagctcgggcacaaggagactcttgcaaggtcttcctcctcgtatcataagctg 60
 QY 1172 CCCCCCTCTGTCATCATGATCAGACCTATGATTTGGCAAAAGCTTCTTCAGAGC 1231
 Db 1172 CCCCCCTCTGTCATCATGATCAGACCTATGATTTGGCAAAAGCTTCTTCAGAGC 1231
 Db 61 cccctctctgcatcatgatacagacattgattctgcaaaagctctctccagagc 120
 QY 1232 TGAACCAAGACCGGCTTCTGGGCGGCAAGGAGGAGGCAAGACCCGCTCTC 1291
 Db 121 tgaaccagagacggtctctggtcggtgaaaggagagagagaccccgctctc 180
 QY 1292 CCACGGATGGGAGAGGCGCAGAGAGACCCAGCCAGTGCCTTTCTTCAGACTGAG 1351
 Db 181 ccacggaatgggagagggcagagagagagccagccaagtgccttctcagactgag 240
 QY 1352 GAGGGGCTTGTTCCTTCCCTCCCTCCCGGACAGCTCCAGGGAGGCTGCTCTGG 1411
 Db 241 gaggggggtgttctccctccctcccgagacaagctccagggcagggctgtccctc 300

```

OY 1412 GCGGCCAGCACTTCCTCAGACACAACTTCTTCTGCTCAGTCGTGGGATCATCA 1471
    |||||||
Db 301 gggccagcagcttcctcagacacacttctctgctgcgcgcgctggtggatcatca 360
OY 1472 CTTACCCACCCCCCAAGTTCAAGACAAATCTTCAGCTGCCCTTCGTCTTCCCTGT 1531
    |||||||
Db 361 ctaccaccccccaagctcagaacaaatcttcagctgccccctcggtgttcccggt 420
OY 1532 GTTGTCTGTAGCTGGGAGCTGCTCCAGACACCAAGAGCCCTCAGCTGTGTAGTCTCC 1591
    |||||||
Db 421 gttgtctgtagctggagctgctcaggaacaaagccctcagctgtgtagctcc 480
OY 1592 CTGACCCCTGTAAATCTCCTTAAGTCTAAGATGATGACTTCAGAAAAA 1643
    |||||||
Db 481 ctgacctgttaattccttaagctaaagatgtagaacttcaaaaaaaaaa 532

RESULT 6
ID AAH02787 standard; cDNA: 537 BP.
AC AAH02787;
XX
XX
XX 14-JUN-2001 (first entry)
DE
XX
XX Prostate tumour antigen cDNA sequence for CGI-69.
KW
XX Human; prostate tumour antigen; prostate tumour; therapy; diagnosis;
KW prostate cancer; immunogenic; cytosolic; vaccine; ss.
OS
XX Homo sapiens.
XX
XX WO200125272-A2.
XX
XX 12-APR-2001.
XX
XX 04-OCT-2000; 2000WO-US27464.
XX
XX 04-OCT-1999; 99US-0157455.
XX
XX (CORI-) CORIAX CORP.
XX
XX
XX Xu J, Skeiky YAM, Reed SG, Cheever MA:
XX
XX WPI: 2001-245062/25.
XX
XX
XX Prostate specific protein and its encoding polynucleotide, useful for
XX the treatment and diagnosis of prostate cancer -
XX
XX Claim 50; Page 244; 276pp; English.
XX
XX The present invention describes an isolated polypeptide (I) comprising
XX at least an immunogenic portion of a prostate tumour antigen protein or
XX its variant. (I) have cytostatic activity and can be used in vaccine
XX production. (I), prostate tumour antigen polynucleotides, an antigen
XX presenting cell (APC e.g. a dendritic cell) that expresses (I), and a
XX pharmaceutical composition containing (I) are useful for inhibiting the
XX development of cancer in a patient. Antibodies specific for prostate
XX specific proteins and oligonucleotides that hybridize to a
XX polynucleotide that encodes a prostate specific protein are useful
XX for detecting the presence or absence of a cancer or monitoring the
XX progression the progression of a cancer, especially prostate cancer.
XX CC AAH02422 to AAH2872, AAH74798 to AAH74821 and AAH74830 are sequences
XX used in the exemplification of the present invention.
XX
XX Sequence 537 BP; 116 A; 167 C; 138 G; 116 T; 0 other;
SQ

```

```

Query Match 32.4%; Score 532; DB 22; Length 537;
Best Local Similarity 100.0%; Pred. No. 6.7e-125;
Matches 532; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
OY 1112 GGGCGAGTCGGGCGACCAAGGACTCTTGGACGCTTCTCCTCGATCATCAAGGCTG 1171
```

```

Db 1 gggccagctggggaccagaagactcttgcaagcttcctccgcgatacataagctg 60
    |||||||
OY 1172 CCCCCTCTGTGGCATATGATATGACACACTATAGTTCGGCAAAAGCTTCTCCAGAGGC 1231
    |||||||
Db 61 cccctctctgtgcatatgataatgacacactatgagcttcggaagcttcccaagagc 120
OY 1232 TGAACAGAGACCGGCTTCTGGCGGCTGAAAGGGGCAAGAGGACGACCGCTCTCTC 1291
    |||||||
Db 121 tgaacacggagccgctctctggtggcgctgaaaggggcaaggaagagaccccgctc 180
OY 1292 CCACGATGGGAGAGGAGGAGAGAGACCCAGGCTGCTTCTTCTTCTTCTTCTTCTTCT 1351
    |||||||
Db 181 ccacgagatgggagagagagagagagagacccagcagcttcttccagacagag 240
OY 1352 GAGGGGGGCTTCTTCCCTCCTCCCTCCGAGACACTTCAGAGGAGGAGGCTGCTCTGG 1411
    |||||||
Db 241 gagggggctgttctccctcctccctcccgagcaagctccagggcaggtctcctctg 300
OY 1412 GCGGCCAGCACTTCCTCAGACACAACTTCTTCTGCTCAGTCGTGGGATCATCA 1471
    |||||||
Db 301 gggccagcagcttcctcagacacacttctctgctgcagctcggtggatcatca 360
OY 1472 CTTACCCACCCCCCAAGTTCAAGACAAATCTTCAGCTGCCCTTCGTCTTCCCTGT 1531
    |||||||
Db 361 ctaccaccccccaagctcagaacaaatcttcagctgccccctcggtgttcccggt 420
OY 1532 GTTGTCTGTAGCTGGGAGCTGCTCCAGACACCAAGAGCCCTCAGCTGTGTAGTCTCC 1591
    |||||||
Db 421 gttgtctgtagctggagctgctcaggaacaaagccctcagctgtgtagctcc 480
OY 1592 CTGACCCCTGTAAATCTCCTTAAGTCTAAGATGATGAACTTCAGAAAAA 1643
    |||||||
Db 481 ctgacctgttaattccttaagctaaagatgtagaacttcaaaaaaaaaa 532

RESULT 7
ID AAC76691/c
ID AAC76691 standard; cDNA: 6712 BP.
XX
XX AAC76691;
XX
XX
XX 08-FEB-2001 (first entry)
DE
XX
XX Human ORFX ORF2246 polynucleotide sequence SEQ ID NO:4491.
XX
XX
XX Human; open reading frame; ORFX; detection; cytosolic; hepatotropic;
XX vunerary; antiproliferative; antiparkinsonian; neurotrophic; neuroprotective;
XX anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
XX immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
XX hypotensive; dermatological; immunosuppressive; antineoplastic; antihypertensive;
XX antiviral; antibacterial; antifungal; antineuritic; antihypertensive;
XX antianemic; gene therapy; cancer; proliferative disorder; hypertension;
XX neurodegenerative disorder; osteoarthritis; graft vs host disease;
XX cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
XX cholesterol ester storage; systemic lupus erythematosus; infection;
XX severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
XX allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
XX bone damage; cartilage damage; antiinflammatory disease; coagulation;
XX thrombosis; contraceptive; ss.
XX
XX
XX Homo sapiens.
XX
XX WO200058473-A2.
XX
XX
XX 05-OCT-2000.
XX
XX
XX 31-MAR-2000; 2000WO-US08621.
XX
XX
XX 31-MAR-1999; 99US-0127607.
XX
XX 02-APR-1999; 99US-0127636.
XX
XX 05-APR-1999; 99US-0127728.
XX
XX 30-MAR-2000; 2000US-0540763.
XX

```

XX (CURA-) CURAGEN CORP.
 PA Shinkets RA, Leach M;
 PI WPI: 2000-602362/57.
 XX P-PSDB: AAB42482.
 DR
 XX Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 PS
 PS Claim 5: Page 3671-3675; 5507pp; English.
 XX
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
 CC antiproliferative; antiparasitism; neurotrophic; vasoregulatory;
 CC osteoplastic; anticonvulsant; antiallergic; immunosuppressive;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasoregulatory;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antiinflammatory; antibacterial; antiviral; antifungal; antihemetic;
 CC antihypertensive; antihemorrhagic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 CC
 XX Sequence 6712 BP: 2319 A; 1206 C; 1244 G; 1942 T; 1 other;
 XX
 SQ

Query Match 14.8%; Score 243.8; DB 21; Length 6712;
 Best Local Similarity 56.3%; Pred. No. 9.2e-52;
 Matches 477; Conservative 0; Mismatches 367; Indels 3; Gaps 1;

QY 389 CCAATGGAAGTCCCTCTGATTCATGATGCTGAGCCCTGTCATGTCGCCCA 448
 Db 3568 CCAAGGAAAGTCTTGTATGATGATGATGATGATGATGATGATGATGATG 3509
 QY 449 ATGTGCGCCGCTGCTCCACCTGCTTCAAGACCCCTACCCGCTTCACTGACCATGATG 508
 Db 3508 AGGGAG---GCMAACAACATATGATTAAGAGCCAGAAATTTCCAGGGAACATTGATG 3452
 QY 509 CCTTGTGAAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 568
 Db 3451 CATTGTTTAAATCAATTCGAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3392
 QY 569 CCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 628
 Db 3391 CCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3332
 QY 629 CCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 688
 Db 3331 CTTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3272
 QY 689 TGGCCCGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 748
 Db 3271 TAGCCAGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3212
 QY 749 AGGCTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 808
 Db 3211 AGTCCAGAGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3152
 QY 809 GTGCTGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 868

Db 3151 ATGTTGATTTCCCTTTGGAGGGGCTGGGCTCTACTGTTCTTGAAGATGATCTTCT 3092
 QY 869 CAGCCCTGATGTTTCAATATGAGCTGTAAGAGCTGCTCAATGGGCTCAGGCCGA 928
 Db 3091 CAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3032
 QY 929 AGCAGACACTTCTGTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 988
 Db 3031 TATATGAGCAACATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2972
 QY 988 CAGTCTGATCTTACCTTTGAGCTGTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1048
 Db 2971 CTTTGGCAACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2912
 QY 1049 TGGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1108
 Db 2911 ATGAAAGATCAATTAATTTCTATGCTTTCATGATGATGATGATGATGATGATGATGATG 2852
 QY 1109 TCGGGGCGAGTCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1168
 Db 2851 TTGTTGCTAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2792
 QY 1169 CTGCCCCCTCTGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1228
 Db 2791 TTGCTCTGCTGCTGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2732
 QY 1229 GGCTGAA 1235
 Db 2731 AACAAAA 2725

RESULT 8
 AAT26018
 ID AAT26018 standard; cDNA to mRNA; 174 BP.
 AC AAT26018;
 DT 16-OCT-1996 (first entry)
 XX
 DE Human gene signature HUMGS08254.
 KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
 KW human; cloning; mapping; non-biased library; diagnosis; detection;
 KW cell typing; abnormal cell function; ss.
 OS Homo sapiens.
 PN W09514772-A1.
 XX
 PD 01-JUN-1995.
 XX
 PE 11-NOV-1994; 94WO-JP01916.
 XX
 PR 12-NOV-1993; 93JP-0355504.
 PA (MATS/) MATSUBARA K.
 XX (OKUBO/) OKUBO K.
 PI Matsubara K, Okubo K;
 DR WPI: 1995-206931/27.
 XX
 PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
 PT for diagnosis of abnormal cell function, by preparing cDNA that
 PT reflects relative abundance of corresp. mRNA in specific human
 PT tissues
 XX
 PS Claim 1: Page 1982; 2245pp; Japanese.
 XX
 CC A single-stranded DNA (or its complementary strand or the corresp.
 CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
 CC given in AAT19001-T26837 and which is able to hybridise to part of
 CC human genomic DNA, cDNA or mRNA is claimed. The GS (gene signature)

DR	WPI: 1999-562047/47.
XX	
PT	New HKNG1 polynucleotides useful in diagnosis and treatment of
PT	neuropsychiatric disorders, e.g. bipolar affective disorders and
PT	schizophrenia
XX	
PS	Claim 2; Fig 3A-R; 205pp; English.
XX	
CC	The present sequence represents HKNG1 (Hong Kong new gene 1). HKNG1
CC	is a gene associated with bipolar affective disorder (BAD). HKNG1
CC	polynucleotides are useful to identify compounds modulating HKNG1 gene
CC	expression or HKNG1 polypeptide expression/activity. Compounds inhibiting
CC	or enhancing HKNG1 gene expression or activity in individuals can then
CC	be administered therapeutically to treat HKNG1-mediated disorders,
CC	especially neuropsychiatric disorders e.g. BAD, schizophrenia, or
CC	HKNG1-mediated myopia disorders, such as early-onset autosomal
CC	dominant myopia. The polynucleotides can be used in gene therapy
CC	techniques to treat such disorders. They are also useful in diagnosis
CC	to identify individuals having, or at risk of developing, HKNG1-mediated
CC	disorders due to mutations in the HKNG1 gene. Such mutations especially
CC	result in the production of a protein with a different sequence to
CC	the human full-length HKNG1 polypeptide or splice variant sequences,
CC	especially the substitution of a lysine for a glutamic acid at residue
CC	202 or 184. The polynucleotides are also useful in gene mapping, to
CC	produce probes or primers to identify similar sequences (e.g. mutants
CC	or sequences from different species) and to produce transgenic
CC	animals.
XX	
SQ	Sequence 72604 BP; 20579 A; 15146 C; 14859 G; 21900 T; 120 other;
	Query Match 4.0%; Score 65.4; DB 20; Length 72604;
	Best Local Similarity 82.6%; Pred. No. 2.3e-06;
	Matches 100; Conservative 1; Mismatches 12; Indels 8; Gaps 2;
OY	2 GGAGGCTGAGCAAGAACCGCTTGAAC---GGGGGGGAGTGTGACGTAGCCCAAGT 58
DB	9317 GGAGCCTTAGCAGGAGAGAAATCGCTTAAACCTGGAGGTGGAGCTTGACGTAGCCCAAGT 9258
OY	59 GGTGGCACTGCACCTCCAGCCTGGCAACAGTGGCGAGCGAGAGCCCGCTCTCAAAACAAA 118
DB	9257 CGTGGCACTGCACCTCCAGCCTGG-----GTGAGCAGAGGAGAGCATCTCAAAAAMAA 9203
OY	119 A 119
DB	9202 A 9202
	RESULT 13
ID	AAC77105/c
XX	AAC77105 standard; cDNA: 4231 BP.
AC	
XX	AAC77105;
XX	
XX	08-FEB-2001 (first entry)
DE	
XX	Human ORF72660 polynucleotide sequence SEQ ID NO:5319.
XX	
KW	Human; open reading frame; ORF; detection; cytosstatic; hepatotropic;
KW	vulnerrary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW	anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW	immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW	hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW	antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW	antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW	neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW	cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW	cholesterol ester storage; systemic lupus erythematosus; infection;
KW	severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW	allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW	bone damage; cartilage damage; antinflammatory disease; coagulation;
KW	thrombosis; contraceptive; ss.
XX	

OS	Homo sapiens.
XX	
PM	WO200058473-A2.
PD	
PX	05-OCT-2000.
PF	
XX	
PR	31-MAR-2000; 2000MO-US08621.
XX	
PR	31-MAR-1999; 99US-0127607.
XX	
PR	02-APR-1999; 99US-0127636.
PR	05-APR-1999; 99US-0127728.
PR	30-MAR-2000; 2000US-0540763.
XX	
PA	(CURA-) CURAGEN CORP.
PI	
XX	
PI	Shinkels RA, Leach M;
XX	
DR	WPI; 2000-602362/57.
XX	
DR	P-PSDB; AAB42896.
XX	
PT	Novel nucleic acids and peptides derived from open reading frame X,
PT	useful for treating e.g. cancers, proliferative disorders,
PT	neurodegenerative disorders and cardiovascular disease -
XX	
PX	Claim 5; Page 4490-4493; 5507pp; English.
XX	
XX	AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC	which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC	sequences have activities such as: cytostatic; hepatotropic; vulneryary;
CC	antiproliferic; antiparkinsonian; nootropic; neuroprotective;
CC	osteoprotic; anticoagulant; antiarthritic; immunosuppressant;
CC	immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC	antidiabetic; hypotensive; dermatological; immunosuppressive;
CC	antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC	antithyroid; and antianemic. The sequences can be used for determining
CC	the presence of or predisposition to, or preventing or treating
CC	pathological conditions associated with an ORFX-associated disorder. The
CC	nucleic acids can be used to express ORFX proteins in gene therapy
CC	vectors. The proteins and nucleic acids may be used to treat cancers,
CC	proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC	grat vs host disease, cardiovascular disease, diabetes mellitus,
CC	hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC	erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC	bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC	allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC	nocurnal haemoglobinuria, antiinflammatory disorders; to enhance
CC	coagulation; to inhibit thrombosis; and as a contraceptive.
SQ	
	Sequence 4231 BP; 832 A; 1348 C; 1095 G; 953 T; 3 other;
	Query Match 3.9%; Score 64.2; DB 21; Length 4231;
	Best Local Similarity 82.0%; Pred. No. 1.8e-06;
	Matches 100; Conservative 0; Mismatches 14; Indels 8; Gaps 2.
QY	1 GGGAGGCTGAACAAGAACCGCTTGAAC---GGGAGTGATGTGTGCAGTAGCCAAGA 57
Db	3131 GGGAGGCTGGGCCAAGAAATTGCTTGAACCTGGAGAGGTGGAGTTGCAGTAGCCAAGA 3072
QY	58 TGGTGCCACTGCATTCACGCTGGCAACGTGGCAGAGCGAGAGCCGCTTCAAAACAA 117
Db	3071 TCGAGGCATATGACACTCACCTGG-----GTGANCAGAGACAGACTCCCTCTCAAACAA 3017
QY	118 AA 119
Db	3016 AA 3015
	RESULT 14
	AAF65290
	ID AAF65290 standard; cDNA; 444 BP.
AC	AAF65290;

Sat Apr 20 10:03:16 2002

us-09-840-787-68.rng

Page 12

Db 53 TTTGGCCACTGCTCCAGCCTGGGCCACGCGCACAGAGCGAGACTCTGTCTC 1

Search completed: April 19, 2002, 21:45:56
Job time: 4270 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2002, 19:42:21 ; Search time 88.57 Seconds
(without alignments)
4201.235 Million cell updates/sec

Title: US-09-840-787-68

Perfect score: 1643

Sequence: 1 GGAGCGCTGACAGACAGAA.....GATGACTTCAAAAA 1643

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 11323899 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued_Patents_NA.*
2: /cgn2_6/ptodata/2/1na/5A.COMB.seq.*
3: /cgn2_6/ptodata/2/1na/5B.COMB.seq.*
4: /cgn2_6/ptodata/2/1na/6A.COMB.seq.*
5: /cgn2_6/ptodata/2/1na/6B.COMB.seq.*
6: /cgn2_6/ptodata/2/1na/backfile1.seq.*

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1643	100.0	1643	2	US-08-933-750C-68
2	1643	100.0	1643	3	US-09-234-613-68
3	62.8	3.8	87350	3	US-08-781-891-79
4	62	3.8	5009	3	US-08-978-741-7
5	62	3.8	5009	4	US-09-333-729A-8
6	61	3.7	997	3	US-09-188-930-23
7	61	3.7	1816	3	US-09-188-930-262
8	60.2	3.7	591	4	US-09-385-982-406
9	60.2	3.7	14636	4	US-09-173-914-6
10	59.6	3.6	330	3	US-09-157-177-114
11	59	3.6	176373	3	US-09-128-155-17
12	58.6	3.6	5138	5	PCT-US96-01314-39
13	58.6	3.6	5138	2	US-08-476-062A-39
14	58.4	3.6	2115	1	US-08-395-800A-7
15	58.4	3.6	19011	1	US-08-310-356-36
16	58.4	3.6	19557	1	PCT-US92-06300-1
17	58.4	3.6	246240	2	US-08-724-394A-20
18	58.4	3.6	246240	2	US-08-724-394A-21
19	58.4	3.6	246240	2	US-08-724-394A-22
20	58	3.5	20303	1	US-08-370-875B-6
21	58	3.5	26764	1	US-08-370-875B-1
22	57.8	3.5	562	4	US-09-385-982-442
23	57.8	3.5	606	4	US-09-385-982-413
24	57.8	3.5	632	4	US-09-385-982-177
25	57.8	3.5	657	4	US-09-385-982-335
26	57.8	3.5	1278	4	US-08-909-965C-4
27	57.8	3.5	2839	3	US-08-468-856B-5

28	57.8	3.5	2839	3	US-08-468-859A-5	Sequence 5, App1
29	57.6	3.5	498	4	US-09-085-199B-36	Sequence 36, App1
30	57.6	3.5	1554	1	US-08-370-975B-10	Sequence 10, App1
31	57.4	3.5	1656	1	US-08-324-465-2	Sequence 2, App1
32	57.4	3.5	1656	2	US-08-465-981-2	Sequence 2, App1
33	57.4	3.5	1656	5	PCT-US93-11915-2	Sequence 2, App1
34	57.4	3.5	1725	1	US-08-324-465-5	Sequence 5, App1
35	57.4	3.5	1725	2	US-08-465-981-5	Sequence 5, App1
36	57.4	3.5	1725	5	PCT-US93-11915-5	Sequence 5, App1
37	57.4	3.5	3804	2	US-08-483-488-5	Sequence 5, App1
38	57.4	3.5	9837	1	US-08-832-883-68	Sequence 68, App1
39	57.4	3.5	9837	2	US-08-832-877-68	Sequence 68, App1
40	57.4	3.5	13158	2	US-08-687-080-105	Sequence 105, App1
41	57.2	3.5	7720	4	US-09-318-448-5	Sequence 5, App1
42	57	3.5	18443	4	US-08-078-284-6	Sequence 6, App1
43	56.8	3.5	1200	4	US-09-018-584A-37	Sequence 37, App1
44	56.8	3.5	5375	3	US-08-757-223-7	Sequence 7, App1
45	56.8	3.5	246240	2	US-08-724-394A-20	Sequence 20, App1

ALIGNMENTS

RESULT 1
US-08-933-750C-68
Sequence 68, Application US/08933750C
Patent No. 5932442
GENERAL INFORMATION:
APPLICANT: Lal, Preethi
APPLICANT: Hillman, Jennifer L.
APPLICANT: Bandman, Olga
APPLICANT: Shah, Purni
APPLICANT: Au-Young, Janice
APPLICANT: Yue, Henry
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
NUMBER OF SEQUENCES: 98
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/933,750C
FILING DATE: September 23, 1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0356 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TEXT:
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 1643 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: SYNOCAT01

CLONE: 724157
US-08-933-750C-68

Query Match 100.0%; Score 1643; DB 2; Length 1643;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1643; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy 1 GGGAGGCTGAAGCAAGAGACCGCTTGAACGGGGGTGATGTGCTGAGCCCAAGATG 60
Db 1 GGGAGGCTGAAGCAAGAGACCGCTTGAACGGGGGTGATGTGCTGAGCCCAAGATG 60
Oy 61 TGCCATGCTGACCTCCAGCTGGCAACAGTGGGAGAGAGAGCGGCTGTCCAAACAAAG 120
Db 61 TGCCATGCTGACCTCCAGCTGGCAACAGTGGGAGAGAGAGCGGCTGTCCAAACAAAG 120
Oy 121 TCTCAGTGGGCCAGCCCTCCCTTCTCCATCCCTGGAGTGTGCTGTTGAGTTT 180
Db 121 TCTCAGTGGGCCAGCCCTCCCTTCTCCATCCCTGGAGTGTGCTGTTGAGTTT 180
Oy 181 GGTTCAGAGACTGAAGCTTCAAGATGGCTACACAGACCTCGGGGATCAGCCCTC 240
Db 181 GGTTCAGAGACTGAAGCTTCAAGATGGCTACACAGACCTCGGGGATCAGCCCTC 240
Oy 241 CAGCAAAATGATGGCTTCAGAGACCGGGGCTGTGTTACTCTCTTCATGACACCCCTG 300
Db 241 CAGCAAAATGATGGCTTCAGAGACCGGGGCTGTGTTACTCTCTTCATGACACCCCTG 300
Oy 301 GACGTGGTGAAGGTTGCGCTGCACTGTCAGCGGCCCTCCATGGCCAGAGCTGATGCT 360
Db 301 GACGTGGTGAAGGTTGCGCTGCACTGTCAGCGGCCCTCCATGGCCAGAGCTGATGCT 360
Oy 361 TCCCTCAGAGTGGAGCTCTCCCTATACCAATGGAAGTGGCTCTGATTTGCAATGGT 420
Db 361 TCCCTCAGAGTGGAGCTCTCCCTATACCAATGGAAGTGGCTCTGATTTGCAATGGT 420
Oy 421 GTCTGAGAGCTCTGTACTGTGCCAAATGTTGCCCGCTGTGCCACCTGGTTTCAAGAC 480
Db 421 GTCTGAGAGCTCTGTACTGTGCCAAATGTTGCCCGCTGTGCCACCTGGTTTCAAGAC 480
Oy 481 CCTACCCGCTTACTGTGACACCATGATGCTTGTGAAGATGTGAGGACAGAGGGCAC 540
Db 481 CCTACCCGCTTACTGTGACACCATGATGCTTGTGAAGATGTGAGGACAGAGGGCAC 540
Oy 541 AGGACCTCTGAGAGGGGCTTCCCGGCAACCTGTGATGACTGTGCCAGCTACCCCATC 600
Db 541 AGGACCTCTGAGAGGGGCTTCCCGGCAACCTGTGATGACTGTGCCAGCTACCCCATC 600
Oy 601 TACTTACTGCTATGACCACTGAAGGCTTCTGTGTGTGTCAGACCTTCTGTAC 660
Db 601 TACTTACTGCTATGACCACTGAAGGCTTCTGTGTGTGTCAGACCTTCTGTAC 660
Oy 661 CTCTACGACCTTATGACCACTGAAGGCTTCTGTGTGTGTCAGACCTTCTGTAC 720
Db 661 CTCTACGACCTTATGACCACTGAAGGCTTCTGTGTGTGTCAGACCTTCTGTAC 720
Oy 721 CCCCTGAGACTTATGCGCAAAAGCTGCAAGCTCAGCATGTGCTGACCGAGCTGGT 780
Db 721 CCCCTGAGACTTATGCGCAAAAGCTGCAAGCTCAGCATGTGCTGACCGAGCTGGT 780
Oy 781 GCTGTGTTGCAAGCTGCAAGCTGCTGAGGCTGAGGCTGCTGCTGAGGCTGGGCT 840
Db 781 GCTGTGTTGCAAGCTGCAAGCTGCTGAGGCTGAGGCTGCTGCTGAGGCTGGGCT 840
Oy 841 CCCAGTGCCTTTCAGATGCTGCTTCTAGCCCTGTACTGTGTTCAACATAGCTGTG 900
Db 841 CCCAGTGCCTTTCAGATGCTGCTTCTAGCCCTGTACTGTGTTCAACATAGCTGTG 900
Oy 901 AAGAGCTGGCTCAATGGGCTCAAGGCGGCAAGAGACCAACATTTCTGTGGCATGAGCTTTG 960
Db 901 AAGAGCTGGCTCAATGGGCTCAAGGCGGCAAGAGACCAACATTTCTGTGGCATGAGCTTTG 960
Oy 961 GCTGGTGCATCTCAGGAGAGGCTGCTGAGTGTGACTTACCTTTGACGTTGAAG 1020

```

```

Db 961 GCTGGTGCATCTCAGGAGAGCTGGCTGCACTGCTGACTTACCTTTGAGCTGTAAG 1020
Oy 1021 ACCCAACCCGAGGTCCTCTCTGGAGCATGAGGCTGTGAAGTGAACCCCTCGATGTG 1080
Db 1021 ACCCAACCCGAGGTCCTCTCTGGAGCATGAGGCTGTGAAGTGAACCCCTCGATGTG 1080
Oy 1081 GACTCCACTGCTGCTGCTGCTGAGAGATCCGGGCGGAGTGGGACCAAGGACTCTTT 1140
Db 1081 GACTCCACTGCTGCTGCTGCTGAGAGATCCGGGCGGAGTGGGACCAAGGACTCTTT 1140
Oy 1141 GCAGGCTTCTCTCTGAGTATCAAGGCTGCCCTCTCTGTGCATCATATGACAC 1200
Db 1141 GCAGGCTTCTCTCTGAGTATCAAGGCTGCCCTCTCTGTGCATCATATGACAC 1200
Oy 1201 TATGATTCGCGCAAAAGCTTCTTCCAGAGGCTGAAACCAAGACCGGCTTGGGGGCTGA 1260
Db 1201 TATGATTCGCGCAAAAGCTTCTTCCAGAGGCTGAAACCAAGACCGGCTTGGGGGCTGA 1260
Oy 1261 AAGGGCAAGAGGCAAGACCCGCTCTCTCCACGATGGGAGAGAGGAGAGAC 1320
Db 1261 AAGGGCAAGAGGCAAGACCCGCTCTCTCCACGATGGGAGAGAGGAGAGAC 1320
Oy 1321 CCAGCAAGTGGCTTCTCTGAGCACTGAGGAGAGGAGGCTTCTCTCCCTCCGCGC 1380
Db 1321 CCAGCAAGTGGCTTCTCTGAGCACTGAGGAGAGGAGGCTTCTCTCCCTCCGCGC 1380
Oy 1381 GACAGCTCCAGAGGAGGAGGCTGCTCTGGGGGCGGAGAGCTCTCTCAGACACACTT 1440
Db 1381 GACAGCTCCAGAGGAGGAGGCTGCTCTGGGGGCGGAGAGCTCTCTCAGACACACTT 1440
Oy 1441 CTCTCTGCTGCTCAGTGTGGGATCATCACTTACCCACCCCAAGTTCAAGACAA 1500
Db 1441 CTCTCTGCTGCTCAGTGTGGGATCATCACTTACCCACCCCAAGTTCAAGACAA 1500
Oy 1501 TCTTCCAGCTGCCCCCTTCTGCTTCCCTGTGTTGCTGATGAGGATGCTCCAGA 1560
Db 1501 TCTTCCAGCTGCCCCCTTCTGCTTCCCTGTGTTGCTGATGAGGATGCTCCAGA 1560
Oy 1561 ACCAAGAGCCCTCAGCTGCTGATGCTCCCTGACCTTGTAAATTCCTAAGTCTAA 1620
Db 1561 ACCAAGAGCCCTCAGCTGCTGATGCTCCCTGACCTTGTAAATTCCTAAGTCTAA 1620
Oy 1621 GATGATGAACCTTCAAAAAAAAA 1643
Db 1621 GATGATGAACCTTCAAAAAAAAA 1643

```

RESULT 2

US-09-234-613-68
Sequence 68, Application US/09234613
Patent No. 6132973

GENERAL INFORMATION:

APPLICANT: Lal, Preeti
APPLICANT: Hillman, Jennifer L.
APPLICANT: Bandman, Olga
APPLICANT: Shah, Purvi
APPLICANT: Au-Yang, Janice
APPLICANT: Yue, Henry
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
NUMBER OF SEQUENCES: 98
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/234.613
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/933.750
FILING DATE: September 23, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0356 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 1643 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: SYNDOT01
CLONE: 724157
US-09-234-613-68

Query Match 100.0%; Score 1643; DB 3; Length 1643;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1643; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGAGGCTGAAACAGAGAACCGCTTGAACGGGGGTGGATGTTGACAGTACGCCAAGATG 60
Db 1 GGGAGGCTGAAACAGAGAACCGCTTGAACGGGGGTGGATGTTGACAGTACGCCAAGATG 60
Qy 61 TGCCACTGACTCCAGCCCTGCGACAGTGGCGAGACGAGACGCCGTCTCAAAACAAAAG 120
Db 61 TGCCACTGACTCCAGCCCTGCGACAGTGGCGAGACGAGACGCCGTCTCAAAACAAAAG 120
Qy 121 TCTCAGAGTGGCCAGCCCTCTTCTCGCATCCCTGAGATTGAGTTGAGATT 180
Db 121 TCTCAGAGTGGCCAGCCCTCTTCTCGCATCCCTGAGATTGAGTTGAGATT 180
Qy 181 GGTTCAGAGTGAAGCTTCAAGATGGCTGACAGAGACCCCTGGGGCTACGCCCTC 240
Db 181 GGTTCAGAGTGAAGCTTCAAGATGGCTGACAGAGACCCCTGGGGCTACGCCCTC 240
Qy 241 CAGCAAAATGTTGGCTCAGAGCACGGGGCTGTTTACCTCTCTCTCATGACACCCCTG 300
Db 241 CAGCAAAATGTTGGCTCAGAGCACGGGGCTGTTTACCTCTCTCTCATGACACCCCTG 300
Qy 301 GACGTGTGAAGTTGCGCTGCAATCTCAGCGGCCCTCCATGGCGAGAGCTGATGCT 360
Db 301 GACGTGTGAAGTTGCGCTGCAATCTCAGCGGCCCTCCATGGCGAGAGCTGATGCT 360
Qy 361 TCCCTCAAGCTGTGAAGCTTCCCTATACCAAAATGAAATGGCTCTGTATATGCAATGT 420
Db 361 TCCCTCAAGCTGTGAAGCTTCCCTATACCAAAATGAAATGGCTCTGTATATGCAATGT 420
Qy 421 GTCTGAGAGCTCTTACCTGTGACCAAAATGTTGCGCGCTGTGACACCGTTTCAAGAC 480
Db 421 GTCTGAGAGCTCTTACCTGTGACCAAAATGTTGCGCGCTGTGACACCGTTTCAAGAC 480
Qy 481 CCTACCCCTCTCACTGAGCAGCATGATGCTTGTGAAGATGTAAGGACAGGAGGACAC 540
Db 481 CCTACCCCTCTCACTGAGCAGCATGATGCTTGTGAAGATGTAAGGACAGGAGGACAC 540
Qy 541 AGGACCTCTGAGAGGCGCTCCCGCCACCCCTGTGATGATCTGTCCACAGTACGCCATC 600
Db 541 AGGACCTCTGAGAGGCGCTCCCGCCACCCCTGTGATGATCTGTCCACAGTACGCCATC 600

Qy 601 TACTTCACTGCTATGACCACTGAAGGCTTCTGTGTGTGACAGCCCTGACCTTCAG 660
Db 601 TACTTCACTGCTATGACCACTGAAGGCTTCTGTGTGTGACAGCCCTGACCTTCAG 660
Qy 661 CTCTACGACCCATGTTGGCTGGCGGCTGGCCGCTTGGGACAGCTGATGATCAGC 720
Db 661 CTCTACGACCCATGTTGGCTGGCGGCTGGCCGCTTGGGACAGCTGATGATCAGC 720
Qy 721 CCCCTGAGACTTATGCGGCAAAAGCTGAGGCTCAGCATGTGTCGATGAGGAGCTGG 780
Db 721 CCCCTGAGACTTATGCGGCAAAAGCTGAGGCTCAGCATGTGTCGATGAGGAGCTGG 780
Qy 781 GCCTGTGTGACACTGACATGAGTGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 840
Db 781 GCCTGTGTGACACTGACATGAGTGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 840
Qy 841 CCCACTGCCCTTTCAGAGTGTGCTTCTCAGCCCTGTACTGTCTCACTATGAGCTGTG 900
Db 841 CCCACTGCCCTTTCAGAGTGTGCTTCTCAGCCCTGTACTGTCTCACTATGAGCTGTG 900
Qy 901 AAGAGCTGGCTAATGGGCTCAGGGCGAAGGACCAAGCTTCTGTGGGCTGAGCTTGTG 960
Db 901 AAGAGCTGGCTAATGGGCTCAGGGCGAAGGACCAAGCTTCTGTGGGCTGAGCTTGTG 960
Qy 961 GCTGTGGCTATCAGAGGAGGCTGAGTGGCTGAGTGGCTGAGTGGCTGAGTGGCTGAG 1020
Db 961 GCTGTGGCTATCAGAGGAGGCTGAGTGGCTGAGTGGCTGAGTGGCTGAGTGGCTGAG 1020
Qy 1021 ACCCAAGCCAGGCTGCTTGGGAGCGATGAGGCTGTGAGAGTGAACCCCTGCATGTG 1080
Db 1021 ACCCAAGCCAGGCTGCTTGGGAGCGATGAGGCTGTGAGAGTGAACCCCTGCATGTG 1080
Qy 1081 GACTCCACTGCTGCTGCTGCGAGATCCGGGCGATGCGGACCAAGGAGCTT 1140
Db 1081 GACTCCACTGCTGCTGCTGCGAGATCCGGGCGATGCGGACCAAGGAGCTT 1140
Qy 1141 GCAGGCTTCTCTCTCTGATCATCAAGGCTGCCCCCTCTGTGCATCATGATCAGACAC 1200
Db 1141 GCAGGCTTCTCTCTCTGATCATCAAGGCTGCCCCCTCTGTGCATCATGATCAGACAC 1200
Qy 1201 TATGAGTTGGCAAAAGCTTCTTCAAGAGCTGAAACAGAGACCGGCTTGGCGGCTGA 1260
Db 1201 TATGAGTTGGCAAAAGCTTCTTCAAGAGCTGAAACAGAGACCGGCTTGGCGGCTGA 1260
Qy 1261 AAGGGGCAAGGAGGAGAACCCGCTCTCCACAGGATGGGAGAGGAGAGAGAGAC 1320
Db 1261 AAGGGGCAAGGAGGAGAACCCGCTCTCTCCACAGGATGGGAGAGGAGAGAGAGAC 1320
Qy 1321 CCAGCAAGTGGCTTCTCTCAGACTGAGGAGGAGGCTTCTTCCCTCCCTCCGCGC 1380
Db 1321 CCAGCAAGTGGCTTCTCTCAGACTGAGGAGGAGGCTTCTTCCCTCCCTCCGCGC 1380
Qy 1381 GACAACTTCAGAGGAGGAGGCTTCCCTTGGGCGGCGCCAGCACTTCTCAGACAACTT 1440
Db 1381 GACAACTTCAGAGGAGGAGGCTTCCCTTGGGCGGCGCCAGCACTTCTCAGACAACTT 1440
Qy 1441 CTTCTGCTGCTCAGAGTGGGAGGATCACTTCCACCCCGCCCAAGTTCAAGACCAA 1500
Db 1441 CTTCTGCTGCTCAGAGTGGGAGGATCACTTCCACCCCGCCCAAGTTCAAGACCAA 1500
Qy 1501 TCTTCAGAGTGGCCCTTCTGTGTTTCCCTGTTTGTGCTAGCTGAGGATGCTCCAGGA 1560
Db 1501 TCTTCAGAGTGGCCCTTCTGTGTTTCCCTGTTTGTGCTAGCTGAGGATGCTCCAGGA 1560
Qy 1561 ACCAAGAGCCCTCAGAGCTGTGATGCTCCTGACCCCTTGAATCCTTAAGCTTAA 1620
Db 1561 ACCAAGAGCCCTCAGAGCTGTGATGCTCCTGACCCCTTGAATCCTTAAGCTTAA 1620
Qy 1621 GATGATGAACCTCAAAAAA 1643
Db 1621 GATGATGAACCTCAAAAAA 1643

```
RESULT 3
US-08-781-891-79
; Sequence 79, Application US/08781891
; Patent No. 6090620
; GENERAL INFORMATION:
; APPLICANT: Fu, Ying-Hui
; APPLICANT: Yu, Chang-En
; APPLICANT: Oshima, Junko
; APPLICANT: Mulligan, John T.
; APPLICANT: Schellenberg, Gerald D.
; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
; TITLE OF INVENTION: WERNER'S SYNDROME
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,891
; FILING DATE: 27-DEC-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: NO. 6090620tendurg Ph.D., Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 240052.419
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 79:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 87350 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-781-891-79

Query Match          3.8%; Score 62.8; DB 3; Length 87350;
Best Local Similarity 75.4%; Pred. No. 1.2e-06;
Matches 92; Conservative 0; Mismatches 27; Indels 3; Gaps 1;

OY 1 GGGAGGCTGAGCAGAGAACCGCTTGAAC---GGGGGTGATGTTGCAGTGACCCAGCA 57
    ||||||| || ||||| ||||||| || || ||||| ||||||| ||
DB 10278 GGGAGGCTGAGCAGAGAAATGCTTGAACCTGGAGGCGGAGGTTGCCGTGAGCCGAGA 10337
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
OY 58 TGGTCCACTGCATCCAGCCCTGGCAGACAGTGCGCAGAGCGGCGTCCAAAACA 117
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
DB 10338 TTGGCCACTGCATCCAGCCCTGGCAGACAGTGCGCAGACACTCTTGTCTCAAAAAAACA 10397
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
OY 118 AA 119
DB 10398 AA 10399

RESULT 4
US-08-978-741-7
; Sequence 7, Application US/08978741
; Patent No. 6100076
; GENERAL INFORMATION:
; APPLICANT: Yang Wang, Michael W. Spellman
; APPLICANT: Yang Wang, Michael W. Spellman
; TITLE OF INVENTION: O-Fucosyltransferase
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
```

```
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Winpatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/978,741
; FILING DATE: 26-NOV-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: 08/792498
; APPLICATION NUMBER: 08/792498
; FILING DATE: 31
; ATTORNEY/AGENT INFORMATION:
; NAME: Svoboda, Craig G.
; REGISTRATION NUMBER: 39,044
; REFERENCE/DOCKET NUMBER: P1041P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1489
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5009 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-978-741-7

Query Match          3.8%; Score 62; DB 3; Length 5009;
Best Local Similarity 73.6%; Pred. No. 6e-07;
Matches 109; Conservative 0; Mismatches 30; Indels 9; Gaps 2;
```

```
OY 1 GGGAGGCTGAGCAGAGAACCGCTTGAAC---GGGGGTGATGTTGCAGTGACCCAGCA 57
    ||||||| || ||||| ||||||| || || ||||| ||||||| ||
DB 4665 GGGAGGCTGAGCAGAGAAATGCTTGAACCTGGAGGCGGAGGTTGCCGTGAGCCGAGA 4724
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
OY 58 TGGTCCACTGCATCCAGCCCTGGCAGACAGTGCGCAGAGCGGCGTCCTCAAAACA 117
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
DB 4725 TTGGCCACTGCATCCAGCCCTGGGCA-----ACAGAGCGAGACTCCATCCAAAAA 4778
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
OY 118 AAGCTCACAGTGGCCCGACGCTTC 145
DB 4779 AGGAGACTTCATGTGCCCAATTTTC 4806

RESULT 5
US-09-333-729A-8
; Sequence 8, Application US/09333729A
; Patent No. 6270987
; GENERAL INFORMATION:
; APPLICANT: Wang, Yang
; APPLICANT: Spellman, Michael W.
; TITLE OF INVENTION: O-Fucosyltransferase
; FILE REFERENCE: P1041P1-Substitute
; CURRENT APPLICATION NUMBER: US/09/333,729A
; CURRENT FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: US 08/798,741
; PRIOR FILING DATE: 1997-11-26
; NUMBER OF SEQ ID NOS: 21
; SEQ ID NO 8
; LENGTH: 5009
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-333-729A-8

Query Match          3.8%; Score 62; DB 4; Length 5009;
Best Local Similarity 73.6%; Pred. No. 6e-07;
Matches 109; Conservative 0; Mismatches 30; Indels 9; Gaps 2;
```



```

RESULT 6
US-09-188-930-23
Sequence 23, Application US/09188930A
Patent No. 6150502
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Onrust, Rene
APPLICANT: Murison, James Greg
TITLE OF INVENTION: Compositions Isolated
FILE REFERENCE: 11000.1011c1
CURRENT APPLICATION NUMBER: US/09/188,930A
CURRENT FILING DATE: 1998-11-09
NUMBER OF SEQ. ID NOS: 348
SOFTWARE: FastSeq for Windows Version 3.0
SEQ. ID NO. 23
LENGTH: 997
TYPE: DNA
ORGANISM: mouse
US-09-188-930-23

```

conservative 0; Mismatches 125; Indels 0; Gaps 0

RESULT 7
US-09-188-930-262
Sequence 263, Application US/09188930A0
Patent No. 6150502
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Oonrust, Rene
APPLICANT: Marlson, James Greg

	Query Match	Best Local Similarity	Score 61;	DB 3;	Length 1816;
	Matches 136;	Conservative	Pred. No. 7.3e-07;	Mismatches 125;	Indels 0; Gaps 0
QY	480	CCTCATCCCGTTTCACATGGCACCATGATGCCTCTGTGAAGATCGTGAAGGCCACGAGGGGCAC	539		
Db	718	ccgcagacaacatctgcatcgttagtggatctcaacagtgttcgaaagaagggaagc	777		
QY	540	CAGAGACCCCTTGAGGGGGCCCCGCCACCTGGTATGACTGTGCTGCCAGCTACGCCAT	599		
Db	778	caagtaactcttcggcgggcaacgagataatgacctcaaatltgcccttgatcgcgcat	837		
QY	600	CTACTTCACCTGCTATGACCAACTAAGGCTTCCTGCTGGTGGTAGACCCCTGACCTCTTA	659		
Db	838	caaatlcatlgcataatgacgatgaagaagcctgtcgttagtgcataagaagacgcttag	897		
QY	660	CCCTCTACGACCCATGAGGGGCTGGGGCCGCCCCTTGEGGACACCGTACTGTATCAG	719		
Db	898	gatccacgaagaagccttgtgcacagcctccttcggaccagcatgtcccbaagtagcatcta	957		
QY	720	CCCCCTGGAGCTTATGCGGAC	740		
Db	958	cccaatlgagagttctgaaagac	978		

Query Match	3.7%	Score 60.2;	DB 4;	Length 591;
Best Local Similarity	80.3%;	Pred. NO. 7.5e-07;		
Matches 98; Conservative	0;	Mismatches 14;	Indels 10;	Gaps 2;

```

OY      1  GGGAGCTGAAGCAGAAACCGCTTGAC---GGGGGTGATGTTCAGTGAAGCCAGA 57
      11  |||||||  |||||||  |||||||  ||  ||  |||||||  |||||||
DB      305  GGNAGGCTGAAGCAGAGAACTCTTGAACCCAGAGAGAGTTCAGTGAAGCCAGA 246
OY      58  TGTGCCACTGCACCTCCAGCGCTGGCAACAGTGGCCAGAGAGCGCGCTCTCAAAACAA 117
      11  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
DB      245  TTGTGCCACTGCACCTCCAGCGCTGGTGA-----CAGAGCGAGACTCGTCTCAAAAAA 193
OY      118  AA 119
      11  |||||||
DB      192  CA 191

RESULT  9
US-09-173-914-6/C
; Sequence 6, Application US/09173914
; Patent No. 6171857
; GENERAL INFORMATION:
; APPLICANT: Hendrickson, Eric
; TITLE OF INVENTION: A No. 6171857el Leucine Zipper, KARP-1 and
; TITLE OF INVENTION: Methods of Regulating DNA Dependent Protein Kinase Activity
; FILE REFERENCE: 80877/7017/HK
; CURRENT FILING DATE: 1998-10-16
; EARLIER FILING DATE: 1997-10-17
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 14636
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-09-173-914-6

Query Match      3.7%; Score 60.2; DB 4; Length 14636;
Best Local Similarity 64.9%; Pred. No. 2.8e-06;
Matches 124; Conservative 0; Mismatches 58; Indels 9; Gaps 2;

OY      1  GGGAGCTGAAGCAGAAACCGCTTGAC---GGGGGTGATGTTCAGTGAAGCCAGA 57
      11  |||||||  |||||||  |||||||  ||  ||  |||||||  |||||||
DB      8143  GGGGCGCTGAAGCAGAGATAGCTTGAAACCGGAGGCGAGGTCAGTGAAGCCGAGA 8084
OY      58  TGTGCCACTGCACCTCCAGCGCTGGCAACAGTGGCCAGAGAGCGCGCTCTCAAAACAA 117
      11  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
DB      8083  TCGGCGCACTGCACCTCCAGCGCTGGCGA-----CAGAGGAGACACCGCTCTCAAAAAA 8030
OY      118  AAGTCACAGTGGCCCGACCTCTCTTCTGTCGATCCCTGGAGTTGGTCTCTTTGAGG 177
      11  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
DB      8029  AAAAAAATTATTTCTTGTGCGCTGCCCGCCAGCACTGTCATGCAATGATTCATATTAGG 7970
OY      178  TTGTGTTTCA 188
      11  |||||||
DB      7969  GGTGCTTTTCA 7959

RESULT  10
US-09-157-177-114/C
; Sequence 114, Application US/09157177
; Patent No. 6090558
; GENERAL INFORMATION:
; APPLICANT: Butler, John M.
; APPLICANT: Li, Jia
; APPLICANT: Monforte, Joseph A.
; APPLICANT: Becker, Christopher H.
; TITLE OF INVENTION: DNA TYPING BY MASS SPECTROMETRY WITH POLYMORPHIC DNA
; TITLE OF INVENTION: REPEAT MARKERS
; FILE REFERENCE: GSTR:017/GETR017P
; CURRENT FILING DATE: 1998-09-18
; EARLIER FILING DATE: 1997-09-19
; NUMBER OF SEQ ID NOS: 135

```

```

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 114
; LENGTH: 330
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-157-177-114

Query Match      3.6%; Score 59.6; DB 3; Length 330;
Best Local Similarity 86.7%; Pred. No. 8.5e-07;
Matches 78; Conservative 0; Mismatches 9; Indels 3; Gaps 1;

OY      1  GGGAGCTGAAGCAGAAACCGCTTGAC---GGGGGTGATGTTCAGTGAAGCCAGA 57
      11  |||||||  |||||||  |||||||  ||  ||  |||||||  |||||||
DB      306  GGGAGCTGAAGCAGAGAAATGCTTGAACCTGGAGGTGAGTTCAGTGAAGCCAGA 247
OY      58  TGTGCCACTGCACCTCCAGCGCTGGCAACAGTGGCCAGAGAGCGCGCTCTCAAAACAA 117
      11  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
DB      246  TCGGCGCACTGCACCTCCAGCGCTGGCGAAG 217

RESULT  11
US-09-128-155-17/C
; Sequence 17, Application US/09128155
; Patent No. 6117654
; GENERAL INFORMATION:
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: 09404/052001
; CURRENT FILING DATE: 1998-08-03
; EARLIER FILING DATE: 1998-08-03
; EARLIER FILING DATE: 1998-07-02
; EARLIER FILING DATE: 1997-08-04
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 176373
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (1)-(176373)
; OTHER INFORMATION: n = A,T,C or G
US-09-128-155-17

Query Match      3.6%; Score 59; DB 3; Length 176373;
Best Local Similarity 79.3%; Pred. No. 1.6e-05;
Matches 96; Conservative 0; Mismatches 20; Indels 5; Gaps 2;

OY      2  GGGAGCTGAAGCAGAAACCGCTTGAC---GGGGGTGATGTTCAGTGAAGCCAGA 58
      11  |||||||  |||||||  |||||||  ||  ||  |||||||  |||||||
DB      105100  GGGAGCTGAAGCAGAGAAATGCTTGAACCCGAGGTGAGAGCTTGGCGTGAAGCCAGAT 105041
OY      59  GGTGCCACTGCACCTCCAGCGCTGGCAACAGTGGCCAGAGAGCGCGCTCTCAAAACAA 118
      11  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
DB      105040  CGGACGACCTGCACCTCCAGCGCTGG--GCAACAGACAGAGCGAGACTCGTCTCAAAAAA 104983
OY      119  A 119
      11  |||||||
DB      104982  A 104982

RESULT  12
PCT-US96-01314-39
; Sequence 39, Application PC/TUS9601314
; GENERAL INFORMATION:
; APPLICANT: M. Amin
; TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
; TITLE OF INVENTION: ANTAGONISTS

```

NUMBER OF SEQUENCES: 78
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 50Z or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/01314
FILING DATE: 30-JAN-96
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/380,167
FILING DATE: 30-JAN-95
ATTORNEY/AGENT INFORMATION:
NAME: John W. Freeman
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 00786/267001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 5137 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PCT-US96-01314-39

Query Match 3.6%; Score 58.6; DB 5; Length 5137;
Best Local Similarity 66.9%; Pred. No. 4.8e-06;
Matches 117; Conservative 0; Mismatches 49; Indels 9; Gaps 2;

QY 2 GGAGGCTGAAGCAGACGCGCTTGAAC---GGGGGTGATGTGACAGTGGCCCAAGT 58
DB 4783 GGAGGTGAGTAGAGAAATTCTTGAACCTGGAGGTGAGAGTGGCCCAAGT 4842
QY 59 GGTCCACTGCACTCCAGCTGGCAACAGTGGCAGAGGAGCGCGCTCTCAAAACAA 118
DB 4843 TGGCGCATGTGACCTCCAGCTGGCA-----ACACAGGAGACCTCGTCTCAAGGAAA 4896
QY 119 AGCTCAGAGTGGCCAGCGCTCTTCTGCGCATCCCTGGAAGTTGTGCTGTTT 173
DB 4897 AATAAATAAAGCGGCGGCGGCGGACATCCCAACCTTGAGAGCTGTCT 4951

RESULT 13
US-08-476-062A-39
Sequence 39, Application US/08476062A
Patent No. 5877375
GENERAL INFORMATION:
APPLICANT: Arnout, M. Amin
TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,062A
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/216,081
FILING DATE: 21-MAR-1994
APPLICATION NUMBER: 07/637,830
FILING DATE: 04-JAN-1991
APPLICATION NUMBER: 07/539,842
FILING DATE: 18-JUN-1990
APPLICATION NUMBER: 07/212,573
FILING DATE: 28-JUN-1988
ATTORNEY/AGENT INFORMATION:
NAME: Freeman, John W.
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 00786/068003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 5138 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 95...3604
US-08-476-062A-39

Query Match 3.6%; Score 58.6; DB 2; Length 5138;
Best Local Similarity 66.9%; Pred. No. 4.8e-06;
Matches 117; Conservative 0; Mismatches 49; Indels 9; Gaps 2;

QY 2 GGAGGCTAACAAGAGACCGCTTGAAC---GGGGGTGATGTGACAGTGGCCCAAGT 58
DB 4784 GGAGGTGAGTAGAGAAATTCTTGAACCTGGAGGTGAGAGTGGCCCAAGT 4843
QY 59 GGTCCACTGCACTCCAGCTGGCAACAGTGGCAGAGGAGCGCGCTCTCAAAACAA 118
DB 4844 TGGCGCATGTGACCTCCAGCTGGCA-----ACACAGGAGACTCGTCTCAAGGAAA 4897
QY 119 AGCTCAGAGTGGCCAGCGCTCTTCTGCGCATCCCTGGAAGTTGTGCTGTTT 173
DB 4898 AATAAATAAAGCGGCGGCGGCGGACATCCCAACCTTGAGAGCTGTCT 4952

RESULT 14
US-08-395-800A-7
Sequence 7, Application US/08395800A
Patent No. 5807732
GENERAL INFORMATION:
APPLICANT: LOWE, JOHN B
TITLE OF INVENTION: SAME
TITLE OF INVENTION: METHOD FOR PRODUCING THE SAME AND A METHOD OF
TITLE OF INVENTION: GENOTYPING A PERSON
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: USA

```

ATTORNEY/AGENT INFORMATION:
NAME: Goodman, Rosanne
REGISTRATION NUMBER: 32,534
REFERENCE/DOCKET NUMBER: A0856
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 454-3817
TELEFAX: (215) 454-3808
INFORMATION FOR SEO ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 19011 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: exon
LOCATION: 1737..1854
OTHER INFORMATION: /number=1
OTHER INFORMATION: /citation= (11)
FEATURE:
NAME/KEY: intron
LOCATION: 1855..2563
OTHER INFORMATION: /number=1
OTHER INFORMATION: /label= intron1
OTHER INFORMATION: /citation= (11)
FEATURE:

```

```

; LOCATION: 2564...2621
; OTHER INFORMATION: /number= 2
; OTHER INFORMATION: /citation= ([1])
; FEATURE:
;

```

```

1 NAME/KEY: intron
2 LOCATION: 2632..4075
3 OTHER INFORMATION: /number= 2
4
5 FEATURE:
6 NAME/KEY: exon
7 LOCATION: 4076..4208
8 OTHER INFORMATION: /number= 3
9 OTHER INFORMATION: /citation= ([1])
10
11 FEATURE:
12 NAME/KEY: intron
13 LOCATION: 4209..6040
14 OTHER INFORMATION: /number= 3

```

```

1 NAME/KEY: 6041..6253
2 LOCATION: 6041..6253
3 OTHER INFORMATION: /number= 4
4
5 FEATURE:
6 NAME/KEY: intron
7 LOCATION: 6253..6801
8 OTHER INFORMATION: /number= 4
9
10 FEATURE:
11 NAME/KEY: exon
12 LOCATION: 6802..6934
13 OTHER INFORMATION: /number= 5
14
15 FEATURE:
16 NAME/KEY: intron
17 LOCATION: 6935..7758
18 OTHER INFORMATION: /number= 5
19
20 FEATURE:
21 NAME/KEY: exon
22 LOCATION: 7759..7856
23 OTHER INFORMATION: /number= 6
24
25 FEATURE:
26 NAME/KEY: intron
27 LOCATION: 7857..9443
28 OTHER INFORMATION: /number= 6
29
30 FEATURE:
31 NAME/KEY: exon
32 LOCATION: 9444..9573
33 OTHER INFORMATION: /number= 7
34
35 FEATURE:
36 NAME/KEY: intron

```

```

1 LOCATION: 17742..18697
2 FEATURE:
3 NAME/KEY: 5'UTR
4 LOCATION: 1737..1775
5 PUBLICATION INFORMATION:
6 AUTHORS: Mungettl, P P
7 AUTHORS: Ruffner, D E
8 AUTHORS: Kuang, W-J
9 AUTHORS: Dennison, O E
10 AUTHORS: Hawkins, J W
11 AUTHORS: Beattie, W G
12 AUTHORS: Dugalezyk, A
13 TITLE: MOLECULAR STRUCTURE OF THE HUMAN ALBUMIN
14 TITLE: GENE IS REVEALED BY NUCLEOTIDE SEQUENCE WITHIN
15 TITLE: 911-122 OF CHROMOSOME 4
16 JOURNAL: J. Biol. Chem.
17 VOLUME: 261
18 PAGES: 6747-6757
19 DATE: 1986
20 RELEVANT RESIDUES IN SEQ ID NO: 36: FROM 1 TO 19011
21
22 OS-08-310-356-36

```

Query Match	Best Local	Similarity	3.68;	Score 58.4;	DB 1;	Length 19011;
Matches	97;	Conservative	0;	Mismatches 16;	Indels 9;	Gaps 2;
QY	1	GGGAGGCTGAAGCAAGAGAACCGCTTGAC---	GGGGGTGGATGTTGCAGTAGCCAGAGA	57		
Db	15419	GGGAGGCTGAAGCAAGAGAAATGGCTTGACCTTGGAGGCGGAGGTTCAGTAGCCGAGCA				15360
QY	58	TGGTGGCACTGCACTCCAGCCCTGGCAACAGATGGCCAGAGAGAGCCGCTCTCAAAACAA				117
Db	15359	TGGCAACCACTGCACTCCAGCCTGGACA-----ACAAAGCAGAACTCGCTTCAAAAAAA				15306
QY	118	AA	119			
Db	15305	AA	15304			

Search completed: April 19, 2002, 21:44:45
Job time: 7344 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2002, 19:38:36 ; Search time 1474.2 Seconds
(without alignments)
11976.200 Million cell updates/sec

Title: US-09-840-787-68

Perfect score: 1643
Sequence: 1 GGGAGCGCTGAAGACAGAGAA.....GATGAACCTTCAAAAAAAAA 1643

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estlin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hic:*
10: gb_est1:*
11: gb_est2:*
12: gb_hic:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_hiv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vit:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1388.8	84.5	1402	12	AF119864 Homo sapi
2	908.8	55.3	963	10	AL519986 AL519986
3	864	52.6	910	10	AL570513 AL570513
4	838.8	51.1	884	10	AL572559 AL572559
5	824.4	50.2	1001	10	AL519987 AL519987
6	823	50.1	921	10	AL529906 AL529906
7	814	49.5	925	10	AL516722 AL516722
8	812.8	49.5	911	10	AL528266 AL528266
9	806.6	49.1	1450	12	AK007934 Mus muscu
10	798.8	48.6	816	11	BG470240 602533758
11	796.4	48.5	885	10	AL514846 AL514846
12	795.8	48.4	940	10	AL523298 AL523298

C 13	790.4	48.1	829	10	AL563104
C 14	789.4	48.0	1371	12	AK007276
C 15	786.6	47.9	795	10	AL571090
C 16	786.2	47.9	862	11	BG470248
C 17	774.4	47.1	901	10	AL545150
C 18	771.4	47.0	891	10	AL522522
C 19	770.2	46.9	1453	12	AK002381
C 20	769.4	46.8	792	11	BG470228
C 21	765.2	46.6	880	10	AL515130
C 22	764	46.5	937	10	AL578185
C 23	763	46.4	917	10	AL544360
C 24	758.6	46.2	1491	12	AK019396
C 25	757.2	46.1	791	10	AL522521
C 26	755	46.0	762	10	AL522849
C 27	753.6	45.9	848	10	BE740464
C 28	750.6	45.7	911	11	BG337246
C 29	743.2	45.2	945	10	AL523299
C 30	739	45.0	778	10	AL574636
C 31	739	45.0	783	10	AL529905
C 32	737.8	44.9	774	11	BG831987
C 33	735.8	44.8	841	11	BI259744
C 34	735.4	44.8	781	10	AI005021
C 35	734.6	44.7	847	11	BI333101
C 36	729.2	44.4	918	11	BG574617
C 37	728.6	44.3	775	10	AL515129
C 38	728.4	44.3	834	10	AL514790
C 39	724.4	44.1	802	11	BG775712
C 40	723.8	44.1	810	10	AI279886
C 41	723.2	44.0	810	11	BG749617
C 42	710.6	43.3	881	11	BI085334
C 43	710.2	43.2	809	11	BI092202
C 44	708.8	43.1	814	10	AA688509
C 45	706.2	43.0	731	10	AL516121

ALIGNMENTS

RESULT 1	AF119864	1402 bp	mrna	HTC	08-MAY-2001
LOCUS	AF119864				
DEFINITION	Homo sapiens	PRO2163	mrna, complete cds.		
ACCESSION	AF119864				
VERSION	AF119864.1	GI:7770164			
KEYWORDS	human.				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
REFERENCE	1 (bases 1 to 1402)				
AUTHORS	Zhang,C., Yu,Y., Zhang,S., Wei,H., Zhang,Y., Zhou,G., Bi,J., Liu,M. and He,F.				
TITLE	Functional prediction of the coding sequences of 79 new genes deduced by analysis of cdna clones from human fetal liver				
JOURNAL	Unpublished				
AUTHORS	2 (bases 1 to 1402)				
TITLE	Direct Submission				
JOURNAL	Submitted (13-JAN-1999) Department of Experimental Hematology, Institute of Radiation Medicine, Beijing Taping Road 27, Beijing, Beijing 100850, P. R. China				
FEATURES	Location/Qualifiers				
source	1..1402				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/clone="FLB8041"				
	/tissue_type="liver"				
	/dev_stage="fetus"				
	4..1017				
	/codon_start=1				
	/product="PRO2163"				
	/protein_id="AAF69618.1"				
CDS					

```

/db_xref="GI:7770165"
/translation="MVASGTGAVTSLFMTPLDVVKVRLQSRPMSALPMSRLMS
LSTYKMKCLILYGVLEPIYICNGARGCAAMPFODPREFETGTMFAFVIVHEGRRLM
SLGPARLYMTVPNATAYFTAYDOLAKATCRALFTSDIAYMVAGALALRSTYVYISL
ELMKRLQAOHVSTRELACVKTAVAGGRSLMLGSPALRQVPSALTFTYIELY
KSMNLGRPRDQTSRELSFVAGISRVVAVLPLPEVYVLRQVVALGAEAVAVNPL
HVDSTWLLLRIRAESGCKGIFAGFLPRITIKAAFPSAIMSTEEFKSPFORLMDRL
LGG"
BASE COUNT      261 a      432 c      401 g      308 t
ORIGIN
Query Match      84.5%; Score 1388.8; DB 12; Length 1402;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1393; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 244 CAAATGTGGCTCAGGACACCGGGGCTGTGTACCTCTCTCTTCATGACACCCCTGAC 303
D 1 CAAATGTGGCTCAGGACACCGGGGCTGTGTACCTCTCTCTTCATGACACCCCTGAC 60
QY 304 GTGGTGAAGGTTGCGCTGACGTCACGAGCGCCCTCCATGGCAGAGAGCTGATGCTTCC 363
D 61 GTGGTGAAGGTTGCGCTGACGTCACGAGCGCCCTCCATGGCAGAGAGCTGATGCTTCC 120
QY 364 TCAGACTGTGGAGCTCTCTCTATACCAATGGAAGTGCTCTCTATTTGCAATGTGTC 423
D 121 TCAGACTGTGGAGCTCTCTCTATACCAATGGAAGTGCTCTCTATTTGCAATGTGTC 180
QY 424 CTGAGAGCTGTACCTGTGCGCAATGTGCGCGCTGTGACCTGCTTCAAGACCT 483
D 181 CTGAGAGCTGTACCTGTGCGCAATGTGCGCGCTGTGACCTGCTTCAAGACCT 240
QY 484 ACCGCTTCACTGACGACCATGATGCTTCTGTAAGATCTGAGGACAGAGGACACAG 543
D 241 ACCGCTTCACTGACGACCATGATGCTTCTGTAAGATCTGAGGACAGAGGACACAG 300
QY 544 ACCCTTGGAGAGCGGCTCCCGCCACCTGCTGATGATGTCACGCTACCCCATCTAC 603
D 301 ACCCTTGGAGAGCGGCTCCCGCCACCTGCTGATGATGTCACGCTACCCCATCTAC 360
QY 604 TTCACTGCTATGACCAACTGAAGGCTTCTGTGTGTGTCAGGCGCTGACCTGACCTC 663
D 361 TTCACTGCTATGACCAACTGAAGGCTTCTGTGTGTGTCAGGCGCTGACCTGACCTC 420
QY 664 TACGACCCATGTGTGCTGCGCGCTGCGCTTGGGACCGTGATCTGATCAGCCCC 723
D 421 TACGACCCATGTGTGCTGCGCGCTGCGCTTGGGACCGTGATCTGATCAGCCCC 480
QY 724 CTGAGAGCTTATGCGGACAAAGCTGAGGCTAGCATGTGTACCGGGAGCTGGGTGCC 783
D 481 CTGAGAGCTTATGCGGACAAAGCTGAGGCTAGCATGTGTACCGGGAGCTGGGTGCC 540
QY 784 TGTGTTCGACATGAGTGGCTCAGGAGTGGCTGCGCTACATGCTGCTGGGCGGCC 843
D 541 TGTGTTCGACATGAGTGGCTCAGGAGTGGCTGCGCTACATGCTGCTGGGCGGCC 600
QY 844 ACTGCCCTTGCAGATGTGCCCTTCTCAGCCCTGTAAGCTGTTCAACTATGAGCTGTGAAG 903
D 601 ACTGCCCTTGCAGATGTGCCCTTCTCAGCCCTGTAAGCTGTTCAACTATGAGCTGTGAAG 660
QY 904 AGTGGCTCATGAGGCTCAGGCGGACCAAGACAGACTTCTGTGGGATAGCTTTGTGCT 963
D 661 AGTGGCTCATGAGGCTCAGGCGGACCAAGACAGACTTCTGTGGGATAGCTTTGTGCT 720
QY 964 GGTGGCATCTCAGGAGCGGTGGCTGACGTGACTTACCTTTGACGTGGTAAAGAC 1023
D 721 GGTGGCATCTCAGGAGCGGTGGCTGACGTGACTTACCTTTGACGTGGTAAAGAC 780
QY 1024 CAAAGCCAGGTGCTCTGAGGAGCGATGAGGCTGTGAGAGTGAACCCCTGATGTGAC 1083
D 781 CAAAGCCAGGTGCTCTGAGGAGCGATGAGGCTGTGAGAGTGAACCCCTGATGTGAC 840
QY 1084 TCACACTGTGCTGCTGCGAGATCCGGGCGGAGTCCGGGACCAAGGACTTTTGCA 1143

```

```

D 841 TCACACTGTGCTGCTGCGAGATCCGGGCGGAGTCCGGGACCAAGGACTTTTGCA 900
QY 1144 GGCCTTCCTCTCGATATATGAAGGCTCCCGCTCCCTGTCATCATATGACACCTAT 1203
D 901 GGCCTTCCTCTCGATATATGAAGGCTCCCGCTCCCTGTCATCATATGACACCTAT 960
QY 1204 GAGTTCGCAAAAGCTTCTTCCAGAGCTGAACAGGACCGGCTTCTGCGGCTGAAG 1263
D 961 GAGTTCGCAAAAGCTTCTTCCAGAGCTGAACAGGACCGGCTTCTGCGGCTGAAG 1020
QY 1264 GGGCAGAGGACCAAGGACCCGCTCTCTCCACGATGGGAGAGGACGAGAGACCA 1323
D 1021 GGGCAGAGGACCAAGGACCCGCTCTCTCCACGATGGGAGAGGACGAGAGACCA 1080
QY 1324 GCGAAGTCCCTTCTTCCAGAGCTGAGGAGGAGGAGGCTTCTCCCTTCCCTCCGCGGAC 1383
D 1081 GCGAAGTCCCTTCTTCCAGAGCTGAGGAGGAGGAGGCTTCTCCCTTCCCTCCGCGGAC 1140
QY 1384 AAGCTCCAGGACAGGAGCTGTCCCTCTGCGGCGCCAGCACTTCTCAGACAACTTCT 1443
D 1141 AAGCTCCAGGACAGGAGCTGTCCCTCTGCGGCGCCAGCACTTCTCAGACAACTTCT 1200
QY 1444 CCTGCTGCTCCAGTGTGGGATATCACTTACCAACCCCAAGTTAAAGCAAAATCT 1503
D 1201 CCTGCTGCTCCAGTGTGGGATATCACTTACCAACCCCAAGTTAAAGCAAAATCT 1260
QY 1504 TCCAGCTGCCCCCTGCTGCTTCCCTGCTGTGTGCTGAGTGGGACATGCTCCAGAAC 1563
D 1261 TCCAGCTGCCCCCTGCTGCTTCCCTGCTGTGTGCTGAGTGGGACATGCTCCAGAAC 1320
QY 1564 AAGAGCCCTCAGCCCTGTGATGCTCCCTGACCCCTGTAAATCTTAAGCTAAAGAT 1623
D 1321 AAGAGCCCTCAGCCCTGTGATGCTCCCTGACCCCTGTAAATCTTAAGCTAAAGAT 1380
QY 1624 GATGACTTCAAAAAAAA 1643
D 1381 GATGAAAAAAA 1400

```

```

RESULT 2
AL519986/c 963 bp mRNA EST 13-FEB-2001
LOCUS
DEFINITION AL519986 LTL_NFL004_NBC2 Homo sapiens cDNA clone CSDB005YF09 3
prime, mRNA sequence.
ACCESSION AL519986
VERSION AL519986.1 GI:12783479
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 963)
AUTHORS Li, W.-B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr.
FEATURES
source
1. 963
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSDB005YF09"
/clone_id="LTL_NFL004_NBC2"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain. Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6

```


vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : filiang@lifetech.com URL : <http://fulllength.invitrogen.com>

BASE COUNT 206 a 288 c 298 g 168 t 3 others
ORIGIN

Query Match 55.3%; Score 908.8; DB 10; Length 963;
Best Local Similarity 98.2%; Pred. No. 4.5e-209;
Matches 947; Conservative 3; Mismatches 11; Indels 3; Gaps 3;

650 TGACCTCTGACCTTACGACCAATGAGGCTGCGGCTGGCCGCTTGGGACCGTGA 709
|||||
663 TGACCTCTGACCTTACGACCAATGAGGCTGCGGCTGGCCGCTTGGGACCGTGA 905
|||||
710 CTGTGATCAGCCCGCTGAGCTTATGCGGACAAAGCTGACAGCTCAGCATGTCTAAC 769
|||||
904 CTGTGATCAGCCCGCTGAGCTTATGCGGACAAAGCTGACAGCTCAGCATGTCTAAC 845
|||||
770 GGGAGCTGGGCTGCTGTGTTGAACTGCAAGGCTCAGAGGCTGCGGCTCAGCTGTG 829
|||||
844 GGGAGCTGGGCTGCTGTGTTGAACTGCAAGGCTGCGGCTCAGCTGTG 785
|||||
830 TGGGCTGGGGCCCGCTGAGCTTATGCGGACAAAGCTGACAGCTCAGCATGTCTAAC 889
|||||
784 TGGGCTG-GGGCGCCACTGCTTCTGAGATGTGCTTCTGAGCTTCTGAGCTTCTGAG 726
|||||
890 ATGAGCTGTGTGAAAGCTGAGCTCAATGAGGCTCAGGCGGACAAAGCTGAGCTTCTG 949
|||||
725 ATGAGCTGTGTGAAAGCTGAGCTCAATGAGGCTCAGGCGGACAAAGCTGAGCTTCTG 666
|||||
950 TGACCTTGTGTGCTGTGTTGAACTGCAAGGCTCAGAGGCTGCGGCTCAGCTGTG 1009
|||||
665 TGACCTTGTGTGCTGTGTTGAACTGCAAGGCTGCGGCTCAGCTGTG 606
|||||
1010 ACGTGTGTAAGAACCAAGGCTGAGCTTGGGAGGAGTGAAGCTGTGAGAGTGAAGC 1069
|||||
605 ACGTGTGTAAGAACCAAGGCTGAGCTTGGGAGGAGTGAAGCTGTGAGAGTGAAGC 546
|||||
1070 CCGTGTGATGTGAGCTTCAACCTGAGCTGCTGCGGAGAGTCCGGGCGGAGTGGGAC 1129
|||||
545 CCGTGTGATGTGAGCTTCAACCTGAGCTGCTGCGGAGAGTCCGGGCGGAGTGGGAC 486
|||||
1130 AGGAGCTTGTGAGGCTTCTTCTCGGATCATCAAGGCTGCCCTCTGCTGCTATCA 1189
|||||
485 AGGAGCTTGTGAGGCTTCTTCTCGGATCATCAAGGCTGCCCTCTGCTGCTATCA 426
|||||
1190 TGATCAGCACTATGAGTGTGGGAAAAGCTTCTCCAGAGGCTGAACCAAGGAGCGGCTTC 1249
|||||
425 TGATCAGCACTATGAGTGTGGGAAAAGCTTCTCCAGAGGCTGAACCAAGGAGCGGCTTC 366
|||||
1250 TGGGCGGCTGAAAAGGGGCAAGGAGGACCCGCTCTCTCCACAGGATGGGAGAGG 1309
|||||
365 TGGGCGGCTGAAAAGGGGCAAGGAGGACCCGCTCTCTCCACAGGATGGGAGAGG 306
|||||
1310 CAGAGAGAGACCCAGCAAGTGTCTTCTCAGCACTGAGAGGAGGGGCTTGTTCCT 1369
|||||
305 CAGAGAGAGACCCAGCAAGTGTCTTCTCAGCACTGAGAGGAGGGGCTTGTTCCT 246
|||||
1370 TCCCTCCGCGGAGCAAGTGTCCAGGAGGAGGCTGTCCCTGAGGAGGCGGCAACACTCTC 1429
|||||
245 TCCCTCCGCGGAGCAAGTGTCCAGGAGGAGGCTGTCCCTGAGGAGGCGGCAACACTCTC 186
|||||
1430 AGACACAACCTTCTGCTGCTCAGTCTGTGGGATCATCACTTACCAACCCCAAGT 1489
|||||
185 AGACACAACCTTCTGCTGCTCAGTCTGTGGGATCATCACTTACCAACCCCAAGT 126
|||||
1490 TCAAGACCAATCTTCCAGCTGCGGCTTCTGCTTCTGCTGTGAGTGTGAGCA 1549
|||||
125 TCAAGACCAATCTTCCAGCTGCGGCTTCTGCTTCTGCTGTG- TTGCTGTAGTGTGAGCA 67
|||||

QY 1550 TGTCTCAGAACCAAGAGCCCTCAGCCTGTGCTAGTCTCCCTGACCTTGTAAATTC 1609
|||||
Db 66 TGTCTCAGAACCAAGAGCCCTCAGCCTGTGCTAGTCTCCCTGACCTTGTAAATTC 7
|||||
QY 1610 TTAAT 1613
|||||
Db 6 TTAAT 3
|||||

RESULT 3
AL570513/c 910 bp mRNA EST 16-FEB-2001
LOCUS AL570513 LTI_NF1006.PL2 Homo sapiens cDNA clone CS0D1018YN07 3
DEFINITION prime, mRNA sequence.
ACCESSION AL570513
VERSION AL570513.1 GI:12926896
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 910)
AUTHORS Li, W.B., Gruber, C., Jesssee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr

FEATURES
source
1..910
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1018YN07"
/clone_id="LTI_NF1006.PL2"
/tissue_type="placenta"
/note="Vector: PCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the PCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : filiang@lifetech.com URL : <http://fulllength.invitrogen.com>

BASE COUNT 197 a 260 c 272 g 156 t 25 others
ORIGIN

Query Match 52.6%; Score 864; DB 10; Length 910;
Best Local Similarity 96.0%; Pred. No. 2.9e-198;
Matches 868; Conservative 24; Mismatches 11; Indels 1; Gaps 1;

QY 626 AGGCTTCTGCTGTGCTGAGCCCTGACCTTGAACCTTACGACCAACCTGAGTGGCG 685
|||||
Db 903 ARGCTTCTGCTGTGCTGAGCCCTGACTTGTGACTTACCAACCAATGTGTGGCG 844
|||||
QY 686 CGCTGGCCGCTTGGGACCGCTGACTGTGATCAGCCCTGAGCTTATGCGGACAAAG 745
|||||
Db 843 CGCTGGCCGCTTGGGACCGCTGACTGTGATCAGCCCTGAGCTTATGCGGACAAAG 784
|||||
QY 746 TGCAGGCTCAGCATGTGTGCTACCGGAGAGTGTGCTGTGTTGAACTCAGAGGCTC 805
|||||
Db 783 TGCAGGCTCAGCATGTGTGCTACCGGAGAGTGTGCTGTGTTGAACTCAGAGGCTC 724
|||||
QY 806 AGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 865
|||||
Db 723 AGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 665
|||||
QY 866 TCTCAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 925
|||||

```

Db 664 TCTCAGCCCTGACTGTTCACTATGAGCTGTGAAGMGTGGCTCAATGAGGCTCAGGC 605
QY 926 CGAAGGACACACTTCTGTGGGAGTGGAGCTTGTGGGATCTTCAGGGAGGCTGG 985
Db 604 CGAAGGACACACTTCTGTGGGAGTGGAGCTTGTGGGATCTTCAGGGAGGCTGG 545
QY 986 CTGCACTGCTACTTCTTACCTTTGACGTGTGAAGACCACAGGCTGCTGGGGAG 1045
Db 544 CTGCACTGCTACTTCTTACCTTTGACGTGTGAAGACCACAGGCTGCTGGGGAG 485
QY 1046 CGATGGAGGCTGTGAGATGAACCCCTGCACTGTGAGCTCCAGCTGCTGCTGGGA 1105
Db 484 CGATGGAGGCTGTGAGATGAACCCCTGCACTGTGAGCTCCAGCTGCTGCTGGGA 425
QY 1106 GGATCCGGGCGGAGTCCGGGACCAAGGAGCTCTTGGAGGCTCTGCTGGATCATCA 1165
Db 424 GGATCCGGGCGGAGTCCGGGACCAAGGAGCTCTTGGAGGCTCTGCTGGATCATCA 365
QY 1166 AGGCTGCCCTCTCTGTGCTCATGATGACACCTATGAGTTGGCAAAAGCTTCTTC 1225
Db 364 AGGCTGCCCTCTCTGTGCTCATGATGACACCTATGAGTTGGCAAAAGCTTCTTC 305
QY 1226 AGAGGCTGAACGAGGAGCGGCTTCTGGGGGCTGGAAGGGGCAAGGAGGAGCCCG 1285
Db 304 AGAGGCTGAACGAGGAGCGGCTTCTGGGGGCTGGAAGGGGCAAGGAGGAGCCCG 245
QY 1286 TCTCTCCACGAGATGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1345
Db 244 TCTCTCCACGAGATGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 185
QY 1346 CTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1405
Db 184 CTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 125
QY 1406 CTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1465
Db 124 TTCTGGGGGCGGAGTCTTCTTCAGACACAACTTCTTCTGCTGCTCAAGCTGGGGA 65
QY 1466 TCATCACTTACCAACCCCAAGTTCAGACCAAACTTTCAGAGCTGCCCTTCTGTTT 1525
Db 64 TCATCACTTACCAACCCCAAGTTCAGACCAAACTTTCAGAGCTGCCCTTCTGTTT 5
QY 1526 CCTT 1529
Db 4 CCTT 1

```

RESULT 4
 AL572559/c 884 bp mRNA EST 16-FEB-2001
 LOCUS AL572559 LTI_NFL006_PL2 Homo sapiens cDNA CS0D1008YD05 3
 DEFINITION prime, mRNA sequence.
 ACCESSION AL572559
 VERSION AL572559.1 GI:12930946
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 884)
 L1, M. B., Gruber, C., Jessee, J. and Polayres, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 CONTACT Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 Location/Qualifiers
 1..884
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0D1008YD05"
 /clone_id="LTI_NFL006_PL2"

```

/lisse_type="placenta"
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com url :
http://fulllength.invitrogen.com"
BASE COUNT 194 a 262 c 268 g 155 t 5 others
ORIGIN
Query Match 51.1%; Score 838.8; DB 10; Length 884;
Best Local Similarity 99.2%; Pred. No. 3.6e-192;
Matches 859; Conservative 4; Mismatches 1; Indels 2; Gaps 2;

```

715 ATCAAGCCCTGAGCTTATGGGACAAAGCTGACAGGCTCAGCATGTGTCAGCGGAG 774
 Db 884 ATCAAGCCCTGAGCTTATGGGACAAAGCTGACAGGCTCAGCATGTGTCAGCGGAG 825
 QY 775 CTGGGCTCTGTGCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACT 834
 Db 824 CTGGGCTCTGTGCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACT 765
 QY 835 TGGGGCCCAAGCCCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACT 894
 Db 764 T-GGGGCCCAAGCCCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACT 706
 QY 895 CTGTGAAGAGCTGCTCAATGAGGCTCAAGCCGGAAGACCAAGCTTGTGTGGCATGAC 954
 Db 705 CTGTGAAGAGCTGCTCAATGAGGCTCAAGCCGGAAGACCAAGCTTGTGTGGCATGAC 647
 QY 955 TTGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1014
 Db 646 TTGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 587
 QY 1015 GTAAAGACCCCAAGCCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1074
 Db 586 GTAAAGACCCCAAGCCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 527
 QY 1075 CATGTGACTCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1134
 Db 526 CATGTGACTCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 467
 QY 1135 CTCTTGGAGGCTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1194
 Db 466 CTCTTGGAGGCTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 407
 QY 1195 AGCACTTATGAGTTCGGCAAAAGCTTCTTCAGAGCTGTAACAGAGCCGGCTTGTGGGC 1254
 Db 406 AGCACTTATGAGTTCGGCAAAAGCTTCTTCAGAGCTGTAACAGAGCCGGCTTGTGGGC 347
 QY 1255 GGCTGAAGGGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1314
 Db 346 GGCTGAAGGGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 287
 QY 1315 GGAGAGCCAGCAAGTGGCTTCTTCAGACTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1374
 Db 286 GGAGAGCCAGCAAGTGGCTTCTTCAGACTGAGGAGGAGGAGGAGGAGGAGGAGGAG 227
 QY 1375 CCGGCGCAAGAGCTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1434
 Db 226 CCGGCGCAAGAGCTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 167
 QY 1435 CAACCTTCTCTGCTGCTCAGTGTGGGATCATCTTACCCACCCCAAGTTCAG 1494
 Db 166 CAACCTTCTCTGCTGCTCAGTGTGGGATCATCTTACCCACCCCAAGTTCAG 107
 QY 1495 ACCAAATCTTCAGAGTGGCCCTTGTGTGTTTCTGTTGCTTACTGGGATGTCT 1554

Db	106	ACCAATATCTTCAGGTGCCCCCTTGTTTCCATGTGTTTCTGTAGCTGGCATGTCT	47
OY	1555	CCAGGAACCAAGAAGCCTCAGCCCG	1580
Db	46	CCAGGACCAGAAAGCCCTCAGCCTG	21
RESULT	5		
AL519987			
LOCUS	AL519987	1001 bp	mRNA EST 13-FEB-2001
DEFINITION	AL519987 LTI_NFL004_NBC2 Homo sapiens CDNA clone CS0DB005YP09		5
ACCESSION	AL519987		
VERSION	AL519987.1	GI:12783480	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1. (bases 1 to 1001)		
JOURNAL	Ll,N.B., Gruber,C., Jessse,J. and Polayes,D.		
COMMENT	Full-length cDNA libraries and normalization Unpublished (2001) Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 Evry cedex - France Email: seque@genoscope.cns.fr, Web : www.genoscope.cns.fr. location/qualifiers		
FEATURES			
source	1..1001		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="CS0DB005YP09"		
	/clone_1lb="LTI_NFL004_NBC2"		
	/sex="male"		
	/issue_type="neuroblastoma cells"		
	/lab_host="DH10B"		
	/note="Organ: Brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oliGo(dt) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng liang life technologies, a division of invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"		
BASE COUNT	167 a 309 c 315 g 208 t	2 others	
ORIGIN			
Query Match	50.2%; Score 824.4; DB 10; Length 1001;		
Best Local Similarity	98.7%; Pred. No. 1,1e-188;		
Matches	894; Conservative 0; Mismatches 6; Indels 6; Gaps 6;		
OY	187	CAGGACTGACCTTCATGATGCCTGACGAGACCTGGGGGATGACGCCCTCCAGCAA	246
Db	83	CAGGACTAAGCTTAATGATGCTACCAAGACCTCGGGGATGACGCCCTCCAGCAA	142
OY	247	ATGCTGACCTCAGGACCGGGGCTGTGTACCTCTCTTCATGACAACCCCTGAGACGTG	306
Db	143	ATGCTGACCTCAGGACCGGGGCTGTGTACCTCTCTTCATGACAACCCCTGAGACGTG	202
OY	307	GTAAGGTTGCGCTGACATCTCAGCGGGCCGTCATGCGAGGAGCTGATGCTTCGCGC	366
Db	203	GTAAGGTTGCGCTGACATCTCAGCGGGCCGTCATGCGAGGAGCTGATGCTTCGCGC	262
OY	367	AGACTGTGAGACCTCTCTATACCAAAATGGAAGTGCCTCTGTATTGCAATGTGTCTTG	426
Db	263	AGACTGTGAGACCTCTCTATACCAAAATGGAAGTGCCTCTGTATTGCAATGTGTCTTG	322
OY	427	GAGCTCTGTACCTGTGCGCAAAATGATGCGCGCTGTGCGACCTGGTTTCAGACCTTACC	486
Db	323	GAGCTCTGTACCTGTGCGCAAAATGATGCGCGCTGTGCGACCTGGTTTCAGACCTTACC	382

OY	487	CGCTTCATGGACCATGGATGGATGCCCTTCTGTGAAGATCGGAGACACGAGGGACACGAGAC	546
Db	383	CGCTTCATGGACCATGGATGGATGCCCTTCTGTGAAGATCGGAGACACGAGGGACACGAGAC	442
OY	547	CTTGAGACGGGCTTCCCGCCACCCTTGGTGTGATGACTGTGCAGCTACCGCCATCTACTTC	606
Db	443	CTTGAGACGGGCTTCCCGCCACCCTTGGTGTGATGACTGTGCAGCTACCGCCATCTACTTC	502
OY	607	ACTGCCATTGACCAACTAAAGGCTTCTGTGTGTGTGTGAGGCCCTGACCTCTGACCTCTAC	666
Db	503	ACTGCCATTGACCAACTAAAGGCTTCTGTGTGTGTGTGAGGCCCTGACCTCTGACCTCTAC	562
OY	667	GCACCCATGTGGCTGGGGGCGCTGGCCCGCTTGGGGACCGGACTGTGATACAGGCCCTTG	726
Db	563	GCACCCATGTGGCTGGGGGCGCTGGCCCGCTTGGGGACCGGACTGTGATACAGGCCCTTG	621
OY	727	GAGCTTATGCGGACAAAGCTCAGGCTTCAGCATGTGTCGTACCGGGAGCTGGTGCCGT	786
Db	622	GAGCTTATGCGGACAAAGCTCAGGCTTCAGCATGTGTCGTACCGGGAGCTGGTGCCGT	681
OY	787	GTTGCACTGCAGTGTGCTCAAGGTTTGCTGGCGCTCACTGTGTGGCTGGGCGCCCACT	846
Db	682	GTTGCACTGCAGTGTGCTCAAGGTTTGCTGGCGCTCACTGTGTGGCTGGGCGCCCACT	740
OY	847	GCCCTTGCAGATGTGCCCTTCTCAAGCCCTGTACTGTGTCAACTATATAGCTGTGTAAAGC	906
Db	741	GCCCTTGCAGATGTGCCCTTCTCAAGCCCTGTACTGTGTCAACTATATAGCTGTGTAAAGC	800
OY	907	TGGCTCAATGGGCTCAGGCGCCAGAGACCAAGACTTCTGTGGGCATAGAGCTTGTGGCTGT	966
Db	801	TGGCTCAATGGGCTCAGGCGCCAGAGACTTCTGTGGGCATAGAGCTTGTGTGGCTGT	859
OY	967	GGCATCTGAGGAGCGGTGTGCTGCAGTGTGACTCTTACCCTTGTGACGTGTAAAGCCAA	1026
Db	860	GGCATCTGAGGAGCGGTGTGCTGCAGTGTGACTCTTACCCTTGTGACGTGTAAAGC	918
OY	1027	CGCCAGGTGCGCTGTGGGAGCATGTGAGAGTGAAGCAACCCCTGCACTGTGGACTCC	1086
Db	919	CGCCAGGTGCGCTGTGGGAGCATGTGAGAGTGAAGCAACCCCTGCACTGTGGACTCC	976
OY	1087	ACCTGG	1092
Db	977	ACTGG	982
RESULT	6		
LOCUS	AL529906		
DEFINITION	AL529906 l7L1.NFL001_NBC4 Homo sapiens cdna clone CSDD005YL19	5	
ACCESSION	AL529906		
VERSION	AL529906.1	GI:12793399	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Li,W.B., Gruber,C., Jesse,J. and Polayes,D.		
JOURNAL	Full-length cDNA libraries and normalization		
COMMENT	Unpublished (2001)		
	Contact: Genoscope		
	Genoscope - Centre National de Sequencage		
	BP 191 91006 EVRY cedex - France		
	Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr.		
FEATURES	Location/Qualifiers		
source	1..921		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="CSDD005YL19"		
	/clone_1ib="L7L1.NFL001_NBC4"		
	/sex="male"		

```

/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(17) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifestech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT      154 a      292 c      269 g      193 t      13 others
ORIGIN

Query Match      50.1%; Score 823; DB 10; Length 921.
Best Local Similarity 98.1%; Pred. No. 2,4e-188;
Matches 831; Conservative 10; Mismatches 5; Indels 1; Gaps 1;

OY 187 CAGAGCTGAAGCTTCAGATGCTGACAGGACCTGGGGCATCAGCCCTCCAGCAA 246
    |||||||
Db 76 CAGAGCTGAAGCTTCAGATGCTGACAGGACCTGGGGCATCAGCCCTCCAGCAA 135

OY 247 ATGCTGGCTCAGGACCGGGGCTGTGCTCTCTCTTCATGACACCCCTGAGCTG 306
    |||||||
Db 136 ATGCTGGCTCAGGACCGGGGCTGTGCTCTCTCTTCATGACACCCCTGAGCTG 195

OY 307 GTGAAGGTTCCTGACAGTCTCAGCGGCTCCAGTGGCAGGAGTATGCTCTCC 366
    |||||||
Db 196 GTGAAGGTTCCTGACAGTCTCAGCGGCTCCAGTGGCAGGAGTATGCTCTCC 255

OY 367 AGACTGTGAGACCTCTCTATACCAAAATGAAAGTGCTCTGTATTTGCAATGCTCTG 426
    |||||||
Db 256 AGACTGTGAGACCTCTCTATACCAAAATGAAAGTGCTCTGTATTTGCAATGCTCTG 315

OY 427 GAGCCTCTGTACCTGTGCTCCCAATGATGTCGCCCTGTGCTTCAAGACCTTAC 486
    |||||||
Db 316 GAGCCTCTGTACCTGTGCTCCCAATGATGTCGCCCTGTGCTTCAAGACCTTAC 375

OY 487 CGCTTCACTGGACCATGATGATGCTGTGAAGATCGTAGAGGACGAGGACGACGAGAC 546
    |||||||
Db 376 CGCTTCACTGGACCATGATGATGCTGTGAAGATCGTAGAGGACGAGGACGACGAGAC 435

OY 547 CTCTGAGACGGCTCTCCCGCCACCTGTGATGACTGTGACAGCTACCGCATCTACTTC 606
    |||||||
Db 436 CTCTGAGACGGCTCTCCCGCCACCTGTGATGACTGTGACAGCTACCGCATCTACTTC 495

OY 607 ACTGCTATGACCAACTGAAAGGCTTCTGTGTGTGCTGAGCCCTGACCTTGACCTTAC 666
    |||||||
Db 496 ACTGCTATGACCAACTGAAAGGCTTCTGTGTGTGCTGAGCCCTGACCTTGACCTTAC 555

OY 667 GCACCCATGATGGTGGGGGCGCTGGCCCGCTTGGGACACGATGATGATGAGCCCTG 726
    |||||||
Db 556 GCACCCATGATGGTGGGGGCGCTGGCCCGCTTGGGACACGATGATGATGAGCCCTG 615

OY 727 GAGCTTATGCGGAAAGCTCAGAGCTCAGCATGTGTGTAACCGGAGACTGGTGTCTGT 786
    |||||||
Db 616 GAGCTTATGCGGAAAGCTCAGAGCTCAGCATGTGTGTAACCGGAGACTGGTGTCTGT 675

OY 787 GTTGAAGTCTGAGTGTGCTCAGGGTGGCTGCTACTGTGGCTGTGGCTGGGCCCCACT 846
    |||||||
Db 676 GTTGAAGTCTGAGTGTGCTCAGGGTGGCTGCTACTGTGGCTGTGGCTGGGCCCCACT 735

OY 847 GCCCTTGAGATGTGCTTCTCAGACCTGTACTAGTGTCACTATGAGCTGGTGAAGAGC 906
    |||||||
Db 736 GCCCTTGAGATGTGCTTCTCAGACCTGTACTAGTGTCACTATGAGCTGGTGAAGAGC 795

OY 907 TGGCTCAATGGGCTCAGGCGGAGAGACCACTTCTGTGGCATGAGCTTTGTGGCTGT 966
    |||||||
Db 796 TGGCTCAATGGGCTCAGGCGGAGAGACCACTTCTGTGGCATGAGCTTTGTGGCTGT 854

OY 967 GGCATCTCAGGAGCGGTGTGCACTGTGTAACCTTGTGACGTGTAAAGACCAA 1026

```

```

|||||
Db 855 GGCATCTCAGGAGCGGTGTGCACTGTGTAACCTTGTGACGTGTAAARACCAA 914
OY 1027 CGCCAGG 1033
Db 915 CGCCAGG 921

RESULT 7
AL516722
LOCUS AL516722 925 bp mRNA EST 13-FEB-2001
DEFINITION AL516722 LTI_NFL011_NBC1 Homo sapiens cDNA clone CS0DA007YA10 5
prime, mRNA sequence.
ACCESSION AL516722
VERSION AL516722.1 GI:12780215
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 925)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1..925
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DA007YA10"
/clone_1id="LTI_NFL011_NBC1"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(17) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library is not normalized, but is the control for
the normalized libraries. Library was constructed by Life
Technologies. Contact : Feng Liang Life Technologies, a
division of Invitrogen 9800 Medical Center Drive Rockville
Maryland 20850, USA Fax : (1) 301 610 8371 Email :
fliang@lifestech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT      153 a      294 c      279 g      192 t      7 others
ORIGIN

Query Match      49.5%; Score 814; DB 10; Length 925.
Best Local Similarity 98.5%; Pred. No. 3.5e-186;
Matches 847; Conservative 5; Mismatches 5; Indels 3; Gaps 3;

OY 187 CAGAGCTGAAGCTTCAGATGCTGACAGGACCTGGGGCATCAGCCCTCCAGCAA 246
    |||||||
Db 68 CAGAGCTGAAGCTTCAGATGCTGACAGGACCTGGGGCATCAGCCCTCCAGCAA 127

OY 247 ATGCTGGCTCAGGACCGGGGCTGTGCTCTCTCTTCATGACACCCCTGAGCTG 306
    |||||||
Db 128 ATGCTGGCTCAGGACCGGGGCTGTGCTCTCTCTTCATGACACCCCTGAGCTG 187

OY 307 GTGAAGGTTCCTGACAGTCTCAGCGGCTTCATGGCAGGAGTATGCTCTCTCC 366
    |||||||
Db 188 GTGAAGGTTCCTGACAGTCTCAGCGGCTTCATGGCAGGAGTATGCTCTCTCC 247

OY 367 AGACTGTGAGACCTCTCTATACCAAAATGAAAGTGCTCTGTATTTGCAATGCTCTG 426
    |||||||
Db 248 AGACTGTGAGACCTCTCTATACCAAAATGAAAGTGCTCTGTATTTGCAATGCTCTG 307

OY 427 GAGCCTCTGTACCTGTGCTCCCAATGATGTCGCCCTGTGCTTCAAGACCTTAC 486

```

```

|||||
Db 308 GAGCCTCTGACTGTGGCCAAATGATGCGCGCTGTGGACACTGTTCAAGACCTTACC 367
Oy 487 CGCTTCACTGGCACCATTGATGCTTGTGAAGATCGTAGAGCAGAGGAGCAGGAGACC 546
Db 368 CGCTTCACTGGCACCATTGATGCTTGTGAAGATCGTAGAGCAGAGGAGCAGGAGACC 427
Oy 547 CTCTGAGACGGGCTCCCGCCACCCTGTGATGATGCTGAGCAGCTGACCGCATCTACTTC 606
Db 428 CTCTGAGACGGGCTCCCGCCACCCTGTGATGATGCTGAGCAGCTGACCGCATCTACTTC 487
Oy 607 ACTGCTATGACCAACTGAAGGCTTCTCTGTGTGTGATGAGCCCTGACCTCTGACCTTAC 666
Db 488 ACTGCTATGACCAACTGAAGGCTTCTCTGTGTGTGATGAGCCCTGACCTCTGACCTTAC 547
Oy 667 GCACCCATGTGTGGCTGGCGGCTGGCCGCTTGGGACACCGTGTGATGATGAGCCCTG 726
Db 548 GCACCCATGTGTGGCTGGCGGCTGGCCGCTTGGGACACCGTGTGATGATGAGCCCTG 607
Oy 727 GAGCTTATGGGACAAAGCTGACAGCTGATGCTGATGCTGAGGAGCTGGTGGCTGT 786
Db 608 GAGCTTATGGGACAAAGCTGACAGCTGATGCTGATGCTGATGACCGGAGCTGGTGGCTGT 667
Oy 787 GTTGGAACTGCAATGCTGCAAGGCTGAGGCTGCTCACTGTGCTGGGCTGGGCCCCACT 846
Db 668 KTTGCAACTGCAATGCTGCAAGGCTGAGGCTGCTCACTGTGCTGGGCTGGGCCCCACT 726
Oy 847 GCCCTTGAGATGTCCTTCTGACGCTTCTGACGCTTCTGACGCTTCTGACGCTTCTGACG 906
Db 727 GCCCTTGAGATGTCCTTCTGACGCTTCTGACGCTTCTGACGCTTCTGACGCTTCTGACG 786
Oy 907 TGGCTCAATGGGCTCA-GGCCGAAGGAGACAGACTTCTGTGGGCAATGAGCTTGTGGCTG 965
Db 787 TGGCTCAATGGGCTCAAGGCGCGGAGAGACAGACTTCTGTGGGCAATGAGCTTGTGGCTG 846
Oy 966 TGGCTCAATGGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1025
Db 847 TGGCTCAATGGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 905
Oy 1026 AGCCAGGTCGCTGGGAG 1045
Db 906 AGCCAGGTCGCTGGGAG 925

```

```

RESULT 8
AL528266 911 bp mRNA EST 13-FEB-2001
LOCUS AL528266 LTI_NFL003_NBC3 Homo sapiens cDNA clone CS0DC023YA13 5
DEFINITION prime, mRNA sequence.
ACCESSION AL528266
VERSION AL528266.1 GI:12791759
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 911)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1. 911
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DC023YA13"
/clone_11b="LTI_NFL003_NBC3"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"

```

```

/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a Notti-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 149 a 288 c 275 g 188 t 11 others
ORIGIN
Query Match 49.5%; Score 812.8; DB 10; Length 911;
Best Local Similarity 98.8%; Pred. No. 6.9e-186;
Matches 814; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
Oy 187 CAGGACTGAAGCTTCAAGATGGCTGACAGGACCCCTGGGCAATGAGCCCTCCAGCAA 246
Db 88 CAGGACTGAAGCTTCAAGATGGCTGACAGGACCCCTGGGCAATGAGCCCTCCAGCAA 147
Oy 247 ATGTGGCTTACGACGACCGGGGCTGTGTACCTCTCTTCAATGACACCCCTGGAGCTG 306
Db 148 ATGTGGCTTACGACGACCGGGGCTGTGTACCTCTCTTCAATGACACCCCTGGAGCTG 207
Oy 307 GTGAAGGTTGCTGCTGACGCTGACGCGCCCTCATGAGCAGAGCTGATGCTTCTCTCC 366
Db 208 GTGAAGGTTGCTGCTGACGCTGACGCGCCCTCATGAGCAGAGCTGATGCTTCTCTCC 267
Oy 367 AGACTGTGAGACCTCTCTTATACCAATGAGAGTCCCTCTGATGCAATGTGTCTTG 426
Db 268 AGACTGTGAGACCTCTCTTATACCAATGAGAGTCCCTCTGATGCAATGTGTCTTG 327
Oy 427 GAGCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 486
Db 328 GAGCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 387
Oy 487 CGCTTCACTGGCACCATTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 546
Db 388 CGCTTCACTGGCACCATTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 447
Oy 547 CTCTGAGAGCGCTCCCGCCACCCTGTGTGATGATGCTGCTGCTGCTGCTGCTGCTG 606
Db 448 CTCTGAGAGCGCTCCCGCCACCCTGTGTGATGATGCTGCTGCTGCTGCTGCTGCTG 507
Oy 607 ACTGCTATGACCAACTGAAGGCTTCTGTGTGTGATGAGCCCTGACCTCTGACCTTAC 666
Db 508 ACTGCTATGACCAACTGAAGGCTTCTGTGTGTGATGAGCCCTGACCTCTGACCTTAC 567
Oy 667 GCACCCATGTGTGGCTGGCGGCTGGCCGCTTGGGCAACCGTGTGATGATGAGCCCTG 726
Db 568 GCACCCATGTGTGGCTGGCGGCTGGCCGCTTGGGCAACCGTGTGATGATGAGCCCTG 627
Oy 727 GAGCTTATGGGACAAAGCTGACAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 786
Db 628 GAGCTTATGGGACAAAGCTGACAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 687
Oy 787 GTTGGAACTGCAATGCTGCAAGGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 846
Db 688 GTTGGAACTGCAATGCTGCAAGGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 747
Oy 847 GCCCTTGAGATGTCCTTCTGACGCTTCTGATGCTTCAACTATGAGCTGTGAAGAGC 906
Db 748 GCCCTTGAGATGTCCTTCTGACGCTTCTGATGCTTCAACTATGAGCTGTGAAGAGC 807
Oy 907 TGGCTCAATGGGCTTACGAGCGGAGGAGCAAGCTTCTGTGGGCAATGAGCTTGTGGCTGT 966
Db 808 TGGCTCAATGGGCTTACGAGCGGAGGAGCAAGCTTCTGTGGGCAATGAGCTTGTGGCTGT 867
Oy 967 GGCATCTCAGGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1010
Db 868 GGCATCTCAGGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 911

```

	RESULT	9
AK007934 LOCUS	AK007934 1450 bp mRNA	HTC 05-JUL-2001
DEFINITION	Mus musculus 10 day old male pancreas cDNA, RIKEN full-length enriched library, clone:1810062G21, full insert sequence.	
ACCESSION	AK007934	
VERSION	AK007934.1 GI:12841807	
KEYWORDS	CAP trapper. Mus musculus (strain:C57BL/6J) 10 day old male pancreas CDNA to mRNA, clone_lib=RIKEN full-length enriched mouse CDNA library clone:1810062G21.	
SOURCE		
ORGANISM	Mus musculus	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Mus;	
AUTHORS	1 (bases 1 to 1450)	
TITLE	Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Methods In enzymology. 303, 19-44 (1999)	
JOURNAL MEDLINE PUBMED	992979253 10349636	
REFERENCE	2 (bases 1 to 1450) Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome research. 10 (10), 1617-1630 (2000)	
AUTHORS	3 (bases 1 to 1450) Shibata,K., Itoh,M., Alzawa,K., Nagaoka,S., Sasakura,N., Carninci,P., Konno,H., Akiyama,T., Nishi,K., Katsunaka,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A., Yamanoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujitake,S., Inoue,K., Togawa,K., Ozawa,K., Tanaka,T., Matsui,S., Kawal,J., Koyeda,Y., Ishikawa,T., Yawata,M., Ohara,E., Watabiki,M., Okazaki,Y., Muramatsu,M., Inoue,T., Kita,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multichipillary sequencer Genome research. 10 (11), 1757-1771 (2000)	
JOURNAL MEDLINE PUBMED	20499374 11042159	
REFERENCE	4 (bases 1 to 1450) The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)	
AUTHORS	5 (bases 1 to 1450) Adachi,J., Alzawa,K., Akihira,S., Akimura,T., Anono H., Arai,A., Arikawa,T., Carninci,P., Fukuda,S., Fukushima,I.Y., Furuno.M., Hanagasaki.T., Hara.A., Hayatsu.N., Hiramoto.K., Hirotsugu.T., Horii.F., Imotani,K., Ishii,Y., Itoh,M., Izawa.S., Kato.H., Kawasaki.J., Komai,Y., Konno,H., Kouda.M., Koza.M., Kusuhara.C., Matsuyama.T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki.R., Ono,M., Okazaki,Y., Okiyoda,C., Owat.C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Saeki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sobabe,Y., Suzuki,K., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasuniishi,A., Yoshida,K., Yoshiho,M., Muramatsu,M. and Hayashizaki,Y. Direct Submission	
TITLE	Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome-gsc.riken-go.jp/, Tel:81-45-503-9212,	
JOURNAL	Fax:81-45-503-9216) Please visit our web site (http://genome-gsc.riken-go.jp/) for further details cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research group in Riken Genomic Sciences Center and genome Science Laboratory in RIKEN.	
COMMENT		

Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAAGACAGACAAGATCCAGACTCTTTTITTTTTTTTAA 3'], cDNA was prepared by using trihaloase thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of subtraction to Rot = 20.0. Second strand cDNA was prepared with the primer adapter of sequence 5'-GACGACAGAGGGCGGCCCATTAATTTCGTGCAATTAATAATACCACCCCCCCC 3'. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR.

FEATURES

source	Location/Qualifiers 1..1450
	/organism="Mus musculus" /strain="C57BL/6j" /db_xref="taxon:10090" /db_xref="MGD:MGI:1895562" /db_xref="MGI:MGI:1915316" /clone="181006ZG21" /sex="male" /tissue-type="pancreas" /clone_id="RIKEN full-length enriched mouse cdna library" /dev_stage="10 day old" 98..1177 note="putative" /codon_start=1 /protein_id="BAB25360.1" /db_xref="GI:12841808" /translation="MDPDDPGISPILOQMVASGAVVTSLEMTPLDVYKRVLOSRRSATSELTTPSRFNSLSTKSSSAIQSPKCLIKNGYLEPIKLPNRGKTRWDPDRFTGTLDAIVKIVRHGRTILMSGLPATLVMTVPATAIYPAYDOALFLGOSLSDSLAPYAVGALAIRMGVTVVSPLEVKRLQOHVSRYELASVOAAVYGQSMILMSWGPELRALRDVPSFALYKWEIVELYKSMLSLSLRKDQSYAISFVAGGISGMVAATLTLPEDVTALRODMSGAIVEAVYKPKPRVDSTMLLRIRIRAESGRGLPAFLPRIKAAPSCAIMISTVEFGKSFPORNQEQPLNR"

CDS

	cdna
--	------

BASE COUNT

a	267	c	456	g	402	t	325
---	-----	---	-----	---	-----	---	-----

ORIGIN

```

1   cccttgagattgtgcgtcttttagaggcttggttttagacatgaacctccaagatggcgtga
2   ttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
3   ttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
4   cccttcgttgtggggcgcaagaagaccggctccagcagcatgttaaatccaaagtgaatga
5   105
6   213 ccagagaccttgaggcgcatcacgccctccagcaaaatgtggctcagacgacggggctgt
7   272
8   106 tcagagaccttgaggcgcatcacgccctccagcaaaatgtggctcagacgacggggctgt
9   165
10  273 ggttacctctctcttcattgatgacaccccctggagcgtggtgaaggttcgctgcagctcacg
11  332
12  166 ggctacactccctcttcattgatgacaccccctggatgtggaaggtccgcttcagctcagag
13  225
14  333 gccctccaatggcgacgacgctgatcctctccgcacgactgtagagcctctccattacaacaa
15  392
16  226 acccttcggcacacgacgattgacacactccctccagatttcggaagctctctctacaccaa
17  285
18  333 at-----ggaaagtcctcctgattatgcaatggtgttcctgga
19  428
20  db      attcatctccgctctacagctccccaggaaaagtgctctcttatatgcacatgagatcctgga
21  345
22  429 gcccttgatcctgaccccaaatgtggtccggctgrrccacctggtttcaaacacctgacgg
23  488
24  db      ttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
25  346 gcccttgatcctgaccccaaatgtggtccggctgrrccacctggtttcaaacacctgacgg
26  405
27  489 ctcttcagcacaatgatatgctctctgtaataatgctgagcagacagggcacacagacacct
28  548
29  db      ttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
30  406 gtctactgacacacttgatgctcttttgatgaatattgtgggagatagagacactagAACCT
31  465
32  549 ctggaagcggcctccccgccacccctgctgatatgctgctcacctaccccatctacttac
33  608
34  db      ttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
35  466 gtggaagcggcctccccgccacccctgctgatatgctgctcacctaccccatctacttac
36  525

```


[illegible]

```

|||||
Db 722 TAAGACCAACGCGAGGTGCTGCGAGGATGAGAGCGTGTGAGATGAGAACCCCTG 781
OY 1075 CATGTGACTCCACCTGGCTGCTGCTGCGGAGA 1108
|||||
Db 782 CATGTGACTCCACCTGGCTGCTGCTGCGGAGA 815

RESULT 11
AL514846 885 bp mRNA EST 13-FEB-2001
LOCUS AL514846 LTI.NFL006.PL2 Homo sapiens cDNA clone Cl08B014ZF10 5
DEFINITION prime, mRNA sequence.
ACCESSION AL514846
VERSION AL514846.1 GI:12778339
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 885)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
Source
1. 885
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Cl08B014ZF10"
/tissue_lib="LTI.NFL006.PL2"
/note="Vector: pCMVSPORT 6; Site:1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 146 a 283 c 264 g 190 t 2 others
ORIGIN
Query Match 48.5%; Score 796.4; DB 10; Length 885;
Best Local Similarity 99.4%; Pred. No. 6.2e-182;
Matches 819; Conservative 1; Mismatches 2; Indels 2; Gaps 2;

OY 195 AAGCTTCAGAGTGGTGCAGGAGCCCTGGGGGCTTCAGCCCTCCGACGAATGGTGGC 254
Db 64 AGCTTCAGAGTGGTGCAGGAGCCCTGGGGGCTTCAGCCCTCCGACGAATGGTGGC 123
OY 255 CTCAGGACACGGGGCTGTGTTACCTCTCTTCATGACACCCCTGAGCTGGTGAAGT 314
Db 124 CTCAGGACACGGGGCTGTGTTACCTCTCTTCATGACACCCCTGAGCTGGTGAAGT 183
OY 315 TCGCTGAGTCTCAGGGCCCTCCATGCGCAGGAGCTGATGCTTCTCCAGACTGTG 374
Db 184 TCGCTGAGTCTCAGGGCCCTCCATGCGCAGGAGCTGATGCTTCTCCAGACTGTG 243
OY 375 GAGCTCTCCATATACCAATGGAAGTGGCTCTGATGCAATGGTTCCTGGAGCTCT 434
Db 244 GAGCTCTCCATATACCAATGGAAGTGGCTCTGATGCAATGGTTCCTGGAGCTCT 303
OY 435 GTACCTGTGCCAATGAGTGCCTGTGCGACCTGTGTTCAAGACCCCTACCCGCTTCA 494
Db 304 GTACCTGTGCCAATGAGTGCCTGTGCGACCTGTGTTCAAGACCCCTACCCGCTTCA 363

```

```

OY 495 TGGACCATGATGATCCTTCGTGAAGATGTGAGGACAGGAGGACACAGACCTCTGGAG 554
Db 364 TGGACCATGATGATCCTTCGTGAAGATGTGAGGACAGGAGGACACAGACCTCTGGAG 423
OY 555 CGGCTCCCGCCACACCTGTGATGACTGTGCCAGCTACGCCCATCTACTTACTTGGCTA 614
Db 424 CGGCTCCCGCCACACCTGTGATGACTGTGCCAGCTACGCCCATCTACTTACTTGGCTA 483
OY 615 TGACCAATGAAGGCTTCGTTGGTGGAGGACCTGACCTGACCTGCTGACGACCAT 674
Db 484 TGACCAATGAAGGCTTCGTTGGTGGAGGACCTGACCTGACCTGCTGACGACCAT 543
OY 675 GGTGCTGGCGGCTGGCGCCCTTGGGACCGTGAAGTGTGATGACAGCCCTGAGACTTAT 734
Db 544 GGTGCTGGCGGCTGGCGCCCTTGGGACCGTGAAGTGTGATGACAGCCCTGAGACTTAT 603
OY 735 GCGGCAAAAGTGCAGGCTCAGCATGTGCTTACCGGAGAGTGGGTGCTGTGTCGAC 794
Db 604 GCGGCAAAAGTGCAGGCTCAGCATGTGCTTACCGGAGAGTGGGTGCTGTGTCGAC 663
OY 795 TGCAGTGGCTGAGGCTGGCTGGCGCTCAGTGGCTGGCTGGGAGCCCACTGCTTGC 854
Db 664 TGCAGTGGCTGAGGCTGGCTGGCGCTCAGTGGCTGGCTGGGAGCCCACTGCTTGC 722
OY 855 AGATGTGCCCTTTCAGCCCTGTACTGTGTTCACTATGAGTGTGAAGAGCTGGCTCAA 914
Db 723 AGATGTGCCCTTTCAGCCCTGTACTGTGTTCACTATGAGTGTGAAGAGCTGGCTCAA 782
OY 915 TGGGTGACGGCGGAAGACAGACTTCTGTGGGATGAGCTTGTGGTGGTGGATCTC 974
Db 783 TGGGTGACGGCGGAAGACAGACTTCTGTGGGATGAGCTTGTGGTGGTGGATCTC 841
OY 975 AGGACGAGTGGCTGAGTGTGACTCTACCTTTGACGTGTAA 1018
Db 842 AGGACGAGTGGCTGAGTGTGACTCTACCTTTGACGTGTAA 885

RESULT 12
AL523298 940 bp mRNA EST 13-FEB-2001
LOCUS AL523298 LTI.NFL003.NBC3 Homo sapiens cDNA clone CS0DC001YB08 3
DEFINITION prime, mRNA sequence.
ACCESSION AL523298
VERSION AL523298.1 GI:12786791
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 940)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
Source
1. 940
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DC001YB08"
/tissue_lib="LTI.NFL003.NBC3"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center

```


Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
 8371 Email : filangellifetech.com URL :
 http://fulllength.invitrogen.com"
 BASE COUNT 200 a 284 c 290 g 157 t 9 others
 ORIGIN

Query Match 48.4%; Score 795.8; DB 10; Length 940;
 Best Local Similarity 93.2%; Pred. No. 8,8e-182;
 Matches 868; Conservative 7; Mismatches 49; Indels 7; Gaps 4;

651 GACCTCTACCTCTACGACCAACCAGTGGCTGGCGCTGGCCGCTTGGACACCGTAC 710
 652 GAGCCCTACCTCTACCTACGACCAACCAGTGGCTGGCGCTGGCCGCTTGGACACCGTAC 865
 711 TGTGATCAGCCCCCTGAGCTTATGCGGCAAAAGCTGAGGCTGAGCATGTGTGTACCG 770
 864 TGTGTCACACCCCTGGGCTAT----GCGACAAAGCTGAGGCTGAGCTGTGTGTACCG 809
 771 GAGCTGGGTGCTGTGTGTGCACTGCAAGTGGCTGAGGCTGGCGCTGACTGTGACT 830
 808 GAGCTGGGTGCTGTGTG-TGCACTGMAAGTGGCTGAGGCTG-TGGCGCTGACTGTGACT 751
 831 GGGCGGG 890
 750 GGGCT-GGGCCCACTGCTGCTGAGATGTCCTTCTACACCCCTGAACTGTGTTCACTA 692
 891 TGAGCTGTGAAAGAGCTGCTCAATGGGCTGAGGCGGCAAGGACAGACTTGTGGGAT 950
 691 TGAGCTGTGAAAGAGCTGCTCAATGGGCTGAGGCGGCAAGGACAGACTTGTGGGAT 632
 951 GAGCTTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1010
 631 GAGCTTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 572
 1011 CGTGTAAAGAGCCCAAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1070
 571 CGTGTAAAGAGCCCAAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 512
 1071 CCTGATGTGAGCTCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1130
 511 CCTGATGTGAGCTCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 452
 1131 GGGACTCTTGTGAGCTCTTCTCTCGATCATCAAGGCTGCTGCTGCTGCTGCTGCTG 1190
 451 GGGACTCTTGTGAGCTCTTCTCTCGATCATCAAGGCTGCTGCTGCTGCTGCTGCTG 392
 1191 GATCAGCAGCAGTATGTTGGGCAAAAGCTCTTCCAGAGCTGAAACGAGACGGCTTCT 1250
 391 GATCAGCAGCAGTATGTTGGGCAAAAGCTCTTCCAGAGCTGAAACGAGACGGCTTCT 332
 1251 GGGCGGCTGAAAGGGGCAAGAGGCAAGACCCGCTCTCTCCACGAGTGGGAGAGGGC 1310
 331 GGGCGGCTGAAAGGGGCAAGAGGCAAGACCCGCTCTCTCCACGAGTGGGAGAGGGC 272
 1311 AGGAGAGACCCCAAGAGGCTCTTCTCTCAGACTGAGAGAGAGGGCTTGTTCCTT 1370
 271 AGGAGAGACCCCAAGAGGCTCTTCTCTCAGACTGAGAGAGAGGGCTTGTTCCTT 212
 1371 CCTCCCGGCGAGCAAGCTCCAGGCAAGGCTGCTCTGAGGCGGCGGCGGCGGCGGCT 1430
 211 CCTCCCGGCGAGCAAGCTCCAGGCAAGGCTGCTCTGAGGCGGCGGCGGCGGCGGCT 152
 1431 GACACAACTTCT 1490
 151 GACACAACTTCT 92
 1491 CAAGACCAAACTTCT 1550
 91 CAAGACCAAACTTCT 32
 1551 GTCTCCAGGAACCAAGAGGCTCTGAGCTTG 1581

Db 31 GTCTCCAGGAACCAAGAGGCTCTGAGCTG 1

RESULT 13
 AL563104/c
 LOCUS
 DEFINITION
 AL563104 LTI_NFL003_NBC3 Homo sapiens cDNA clone CS0DC023YA13 3
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE
 1 (bases 1 to 829)
 LA,W.B., Gruber,C., Jesse,J. and Polayes,D.
 Full-length cDNA libraries and normalization
 JOURNAL
 Unpublished (2001)
 CONTACT: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 Location/Qualifiers
 1. 829

FEATURES
 source
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0DC023YA13"
 /clone_lib="LTI_NFL003_NBC3"
 /sex="male"
 /tissue_type="neuroblastoma cells"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with NotI and cloned into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized. Library of the pCMVSPORT 6 by life technologies. Contact: Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : filangellifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 180 a 247 c 241 g 151 t 10 others
 ORIGIN

Query Match 48.1%; Score 790.4; DB 10; Length 829;
 Best Local Similarity 97.0%; Pred. No. 1.7e-180;
 Matches 805; Conservative 8; Mismatches 16; Indels 1; Gaps 1;

747 GCAGGCTCAGCATGTGTCGTAACCGGAGGCTGGCTGCTGTTCGAACCTGCACTGCTCA 806
 829 GCAGGCTCAGCTTCTTTCGTAACCGGAGGCTGGCTGCTGTTCGAACCTGCACTGCTCA 770
 807 GGGTGGCTGGCGCTCACTGTGGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTT 866
 769 GGGTATGTGGCGCTCACTGTGGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTT 711
 867 CTGAGCCCTGTAATGCTTCAATATAGCTGGTGAAGAGCTGCTCAATGGCTCAGGCC 926
 710 CTGAGCCCTGTAATGCTTCAATATAGCTGGTGAAGAGCTGCTCAATGGCTCAGGCC 651
 927 GAAGGACAGACTTCTGTGGGCTGAGCTTGTGGCTGGGCTGAGCTGAGGAGGCTGGC 986
 650 GAAGGACAGACTTCTGTGGGCTGAGCTTGTGGCTGGGCTGAGCTGAGGAGGCTGGC 591
 987 TGCAGTGTGACTTACCTTTGACGTGTGTAAGAGCAACGCAAGTGTGCTTGGAGAC 1046
 590 TGCAGTGTGACTTACCTTTGACGTGTGTAAGAGCAACGCAAGTGTGCTTGGAGAC 531
 1047 GATGAGGCTGTGAGAGTAACCCCTGATGATGAGCTCAGCTGCTGCTGCTGCTGCTG 1106
 530 GATGAGGCTGTGAGAGTAACCCCTGATGATGAGCTCAGCTGCTGCTGCTGCTGCTGCTG 471

[illegible]

MEDLINE	20530913
PUBMED	11076861
REFERENCE	4 (bases 1 to 1371)
AUTHORS	The Riken Genome Exploration Research Group Phase II Team and the FANTOM Consortium
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 685-690 (2001)
AUTHORS	Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Atakawa,T., Carinini,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiroaka,T., Horii,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishii,K., Nomura,K., Numasaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sobabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,A., Tonaka,T., Tsujima,Y., Toyota,T., Yamamura,T., Yasunishi,Y., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE	Direct Submission
JOURNAL	Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT	Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
FEATURES	location/Qualifiers
source	1..1371
	/organism="Mus musculus"
	/strain="C57BL/6J"
	/db_xref="taxon:10090"
	/db_xref="MGD:MGI:1896541"
	/db_xref="MGI:MGI:1915316"
	/clone="1/00125C18"
	/sex="male"
	/tissue-type="testis"
	/clone_idb="RIKEN full-length enriched mouse cDNA library"
	/dev_stage="adult"
	123..1133
CDS	/note="putative"
	/codon_start=1
	/protein_id="BAB24932.1"
	/db_xref="gi:12840722"
	/translation="MDQDDPWGISPLQQMVASGAGAVTSLFMTPLDYVKRVLQSRRP SATSEIITPSFEWLSSTKSSALSPGKCLLNGVLEPLYLCPNGRATFWQDPFL RFETGLDAFKIVTHHEGRITMSGLPTLVIAVPAVLAIFYAYOLAFLCGISLSD LYAPRWGAALARAKTVYVYSLEKLPRKLOAHVSYLEASAOAAVTGGMRSLIMG WGPAIRKDVPSALSALWENTVELKSNLSLRKDQTSGVISPVAGISGMVAATILIPP DVVTFQRORMSLGAVEAVERKKPPRDSTWLFLAESGLINLPADISLQASPSSRLHPVP PS"
BASE COUNT	255 a 437 c 387 g 292 t
ORIGIN	
Query Match	48.0%; Score 789.4; DB 12; Length 1371;
Best Local Similarity	84.0%; Pred. No. 3,3e-180;
Matches 925; Conservative 0; Mismatched 151; Indels 25; Gaps 2;	
187 CAGGACTGAAGCTTCANAGTGCCTACGACGAGACCCTGGCGGGATCATGCCCTCCAGCAA	246
Db 105 CAGGATTGAAATTCCAAGATGATGATCGAGACCTTGGGCATTAAGCCCCCTTAGGCA	164
247 ATGGGGCCCTCAGGACCGGGGGGTGGTACTCTCTTTATGACACCCCTGGACGCG	306
Db 165 ATGGTGGCTTCAGAGACCCGGGGCTGTGGTACCTCCCTTTATATACACCCCTGGATGG	224
307 GTGAAGGTTGCGCTGCAGTCTCAGCGGGCCCTCCATGGCCAGCGAGCTGATGCTTCCTCC	366
Db 225 GTGAAGGTCGCGCTTCAGTCTCAGAGACCCCTGGCAACCGAGGAATTGACAACCTCCCTCC	284

QY	367	AGACGTGGAGCCCTCTCTATACCAAT-----GGAGATGC	407
Db	285	AGATTCTGGATCTCTCTACACCAATATCTCCGCTACATCTCCCGAGGAGTGC	344
QY	403	CTCTGTATTGCATGGTGTCTGAGACCTGTACCTGTGGCCAAATGGTCCCCGTGT	462
Db	345	CTCTTATACGCAATGAGTCCGTGGAGCCCTGTACCTGTGCCAAATGGTACCCTGT	404
QY	463	GGCACCGTGGTTCAAGACCTTACCCGCTTCACTGGCACCATGATGCTTGTGTAATC	522
Db	405	GGCACCTGGTTTGAAGACCCACACCGTGTACCTGGCACCTTGATGCTTGTGAATTT	464
QY	523	GTCAGGCACAGGAGGACACAGGACCCCTGTGGACGGGCCCCCGGACACCTGTGTACT	582
Db	465	GTCGGGCATGAGGGACATAGAACCTGTGTGAGCGGCTCCCAAGTACCCGTGTATGACC	524
QY	583	GTCGCCAGCTACGCCCATCTTACTTCACTGTACGTATGACCAATGAAAGGCTTCTGTGGT	642
Db	525	GTCGCCAGCTACTGTCTATCTACTTCACTGTACTGACCTTACGACCAATCAAGGCTTCTGTGTGT	584
QY	643	CGAGCCCTACACCTGTACCTTACGACACCCATGGTGGCTGGCCGCTGGCCGCTTGGGC	702
Db	585	CAGTCTTACCTCTTACCTCTTACGACACCCATGGTGGCTGGTCCCTGCTCAAAAGGGC	644
QY	703	ACCGTACTGTATACGCCCCCTGGAGGCTTATGCGGACAAAGCTGACAGGCTACGATGTG	762
Db	645	ACCGTACTAGTTGTACAGCCCTTGGAGCTGTGGGACCAAGCTGACAGGCTACGATGTG	704
QY	763	TGCTACCGGAGACTGGGTGCTGTCTTGAACCTCACTAGGCTGAGGTGGCTGGCTCA	822
Db	705	TGCTACCGGAGACTGGGTGCTGTCTTGAACCTCACTAGGCTGAGGTGGCTGGCTCT	764
QY	823	CTGTGGCTGGGTGGGGCCCACTAGCCCTTGAAGATGCTGACCTTACGCTGTATCTGG	882
Db	765	CTGTGGCTGGGTGGGGCCCACTAGCTTTCGAGATGTCCCCCTTTCACACTGTATCTGG	824
QY	883	TTCACCTTACGAGCTGGTGAAGAGTGGCTCAATGGGCTCAGGCCGAAGACACAGATCT	942
Db	825	TTCACCTACGAGTGGTGAAGAGTGGCTCAGGGGACTAGAACCAAAACACAGACATCG	884
QY	943	GTGGGCATGAGCTTGTGGCTGTGGGATCTCAGGACAGGTGGCTGACGTGTGACTTA	1002
Db	885	GTGGGCATGAGCTTGTGGCGGGGATCTCAGGAATGTGGCTTCCACCTTATCCCTTA	944
QY	1003	CCCTTTCAGTGGTGAAGACCAAGCCAGGCTGCTCTGGAGCATGGAGGCTGTGAGA	1062
Db	945	CCCTTTCAGTGGTGAAGACCAAGCAAGATCTCACTGGAGCGGTGAAGGCTGTGAGA	1004
QY	1063	GTCGAACCCCTCGATGTGACTCCACTGGCTGTCTGTGGAAGATCCGGGCGGAGTGC	1122
Db	1005	GTCGAACCCCGCCACAGAGTGGACTCCACTGGCTGTCTGT - TTGCGAAGATCCGGGCTGAATCT	1063
QY	1123	GGCACCAAGGAGACTTTTGGAGGCTTCTTCTCGGATCATCAAGGCTGCCCTCTGT	1182
Db	1064	GGCACCAAGGAGACTTTTGGAGGCTTCTTCTCGGATCATCAAGGCTGCCCTCTGT	1123
QY	1183	GGCATTCATGATCAGCACTATGAGTTGGGCAAAAGCTTTCGAGGCTCAACAGAGAC	1242
Db	1124	GGCATTCATGATCAGCACTATGAGTTTGGCAAAAGCTTTCGAGGCTCAACCAAGAG	1183
QY	1243	CGGCTTGTGGGGCGGTGAAG	1263
Db	1184	CAGGCTGTGGGGCGGTGAAG	1204

RESULT 15	AL571090/c	LOCUS	DEFINITION	AL571090	795 bp	MRNA	EST	16-FEB-2001
ACCESSION	AL571090	LT1.NF006.P12	Homo sapiens cDNA clone	prime, mRNA sequence.				CSDD0128Y09 3
VERSION	AL571090.1	GI:12928041						
KEYWORDS	EST.							

SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eulhelia; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	Li, W. B., Gruber, C., Jesse, J. and Polyes, D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished (2001)
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 101 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES	
SOURCE	Location/Qualifiers 1..795

BASE COUNT	167 a	241 c	231 g	142 t	14 others
ORIGIN					

Query Match	47.98;	Score 786.6;	DB 10;	Length 795;
Best Local Similarity	98.18;	Pred. No. 1.4e-179;		
Matches 780; Conservative	14;	Mismatches 1;	Indels 0.	Cams 0

QY	725	TGGAGCTTATTGCGGACAAAGCTGCAGGCTCAGCATGTGTCTCTACCGGGAGCTGGGTGCTT	784
Db	795	TGGAGCTTATTGCGGACAAAGCTGCAGGCTCAGCATGTGTCTCTACCGGGAGCTGGGTGCTT	736
QY	785	GTTTTCGAAGCTCAGATGGCTCAGAGGTGCTGGAGCTCAATGTGGCTGGGCTGGGGCCCCA	844
Db	735	GTTTTCGAAGCTCAGATGGCTCAGAGGTGCTGGAGCTCAATGTGGCTGGGCTGGGGCCCCA	676
QY	845	GTGGCCCTTCGAGATGTGCCCCCTTTCAGCCCTGTACTGTGGTTCAACTATGAGACTGTGAGA	504
Db	675	GTGGCCCTTCGAGATGTGCCCCCTTTCAGCCCTGTACTGTGGTTCAACTATGAGACTGTGAGA	616
QY	905	GCTGGCTCAATGGGCTCAGGCCGGAAGACACGAACTTGTGGGCATGAGCTTGTGGCTG	964
Db	615	GCTGGCTCAATGGGCTCAGGCCGGAAGACACGAACTTGTGGGCATGAGCTTGTGGCTG	556
QY	965	GTGGCATCTCAGGGAGAGGTGGCTGCTCAATGCTGACTTACCCCTTGTGACGTGTAAAGACC	102
Db	555	GTGGCATCTCAGGGAGAGGTGGCTGCTCAATGCTGACTTACCCCTTGTGACGTGTAAAGACC	496
QY	1025	AACGCCAGGTGGCTCTGTGGAGCGCATGTGAGGCTGTGAGAGTGAACCCCTCGCATGTGGAGCT	108
Db	495	AACGCCAGGTGGCTCTGTGGAGCGCATGTGAGGCTGTGAGAGTGAACCCCTCGCATGTGGAGCT	436
QY	1085	CCACCTGGGCTCTGTGCGGAGGATCCGGGCGCAAGTCGGGACACCAAGGACTCTTTGGAG	1144
Db	435	CCACCTGGGCTCTGTGCGGAGGATCCGGGCGCAAGTCGGGACACCAAGGACTCTTTGGAG	376
QY	1145	GCTTTCCTTCCTCGGATCATCAAGGCTGCCCTCTCTGTCGCATTCATGTATCAGACACTTATG	1204
Db	375	GCTTTCCTTCCTCGGATCATCAAGGCTGCCCTCTCTGTCGCATTCATGTATCAGACACTTATG	316
QY	1205	AGTTGCGCAAAAGCTTCTTCAGAGGCTGAACAGAGACCGGCTTGTGGGCGGCTGAAGAAG	1264
Db	315	AGTTGCGCAAAAGCTTCTTCAGAGGCTGAACAGAGACCGGCTTGTGGGCGGCTGAAGAAG	256

```

QY 1265 GGCAGAGGCGAAGACCCGCTCTCCACGGATGGGAGAGGCGAGAGACCCAG 1324
    |||||||
Db 255 GGCAGAGGCGAAGACCCGCTCTCCACGGATGGGAGAGGCGAGAGACCCAG 196
    |||||||
QY 1325 CCAAGTGCCTTTTCTCAGCACTGAGGAGGAGGAGGCTTTGTTCCCTCCCGGCGACA 1384
    |||||||
Db 195 CMAAGTGCCTTTTCTCAGCACTGAGGAGGAGGAGGCTTTGTTCCCTCCCGGCGACA 136
    |||||||
QY 1385 AGCTCCAGGCGAGGCTGTCCCTCTGGGCGGCGCCAGCACTTCTCAGACAACTTCTTC 1444
    |||||||
Db 135 AGCTCCAGGCGAGGCTGTCCCTCTGGGCGGCGCCAGCACTTCTCAGACAACTTCTTC 76
    |||||||
QY 1445 CTGCTGCTCCAGTGTGGGATCATCACTTACCCACCCCAAGTTCAAGACCAATCTT 1504
    |||||||
Db 75 CTGCTGCTCCAGTGTGGGATCATCACTTACCCACCCCAAGTTCAAGACCAATCTT 16
    |||||||
QY 1505 CCAGCTGCCCCCTTC 1519
    |||||||
Db 15 CCAGCTGCCCCCTTC 1
    |||||||

```

Search completed: April 19, 2002, 21:00:45
 Job time: 4929 sec